# OUTCOMES IN CLINICAL INTERVENTION RESEARCH:

# SELECTION, SPECIFICATION, AND REPORTING

by
Ian Jude Saldanha

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
July, 2015

## DISSERTATION ABSTRACT


### Introduction

Sound clinical intervention research relies on the use of the right outcomes dependably and without bias. This dissertation addresses important research gaps related to selection, specification, data collection and analysis, and reporting of outcomes in research.

### Methods

We examined systematic reviews (SRs) and randomized controlled trials (RCTs). For Aim 1, related to outcome selection, we conducted a case study of outcomes in all Cochrane SRs addressing HIV/AIDS (June 2013), to evaluate whether social network analysis methods could be used to identify central outcomes for core outcome sets. For Aim 2, related to outcome specification and data collection and analysis, we examined all Cochrane SR protocols (June 2013) addressing four major eye conditions for completeness of pre-specification and comparability of outcomes. For Aim 3, related to outcome reporting, we examined all conference abstracts of RCTs presented at the Association for Research in Vision and Ophthalmology (ARVO) 2001-2004 conferences to: (1) evaluate agreement in main outcome results comparing abstracts and corresponding full publications, and (2) evaluate the association between conflicts of interest and full publication.

**Results**

For Aim 1, we applied social network analysis methods to identify *central* outcomes, and found that the most central outcomes often differed from the most *frequent* outcomes. For Aim 2, outcome pre-specification in SRs was largely incomplete. For Aim 3, only 44.8% of abstracts describing RCTs were published in full, and more than half (54.7%) of the 86 conference abstract/full publication pairs had some form of discordance in reported results for the main outcome. First author conflicts of interest were associated with full publication, irrespective of whether the main outcome results in the abstract were statistically significant, not statistically significant, or the statistical significance was not reported.

**Conclusions**

This dissertation identifies causes for concern, such as incomplete outcome pre-specification in SR protocols, and discrepancies in results comparing conference abstract/full publication pairs of the same RCT. We suggest ways forward, such as use of social network analysis methods to identify central outcomes for core outcome sets, and incorporation of a five-element framework for outcome pre-specification in SR and RCT protocols.

## DISSERTATION READERS AND FINAL EXAMINATION COMMITTEE

**Kay Dickersin, PhD**
Dissertation Advisor
Professor
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

**Henry D. Jampel, MD**
Professor
Department of Ophthalmology
Johns Hopkins School of Medicine

**Joanne Katz, ScD**
Chair of Examination Committee
Professor
Department of International Health
Johns Hopkins Bloomberg School of Public Health

**Roberta W. Scherer, PhD**
Senior Scientist
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

## ALTERNATE COMMITTEE MEMBERS

**Marie Diener-West, PhD**
Professor
Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health

**Janet T. Holbrook, PhD**
Associate Professor
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

## ACKNOWLEDGEMENTS

Completing a PhD is a labor of love. One has to love a topic enough to spend the requisite time, money, and efforts to pursue it in depth. For me, arriving at my topic (outcome use in clinical intervention research) involved the confluence of two phases of my academic life: my clinical training and my research training. I have many to thank for helping me realize this love, formulate it into a dissertation, and complete the dissertation successfully.

I am most thankful for my PhD advisor Dr. Kay Dickersin. Words fail to capture how exceptional a mentor, teacher, and role model she is; but, I am going to try. Kay is a perfectionist, and has taught me to never accept mediocrity. She has always been patient and believed in me while I developed and worked through my research ideas. She has always encouraged me to think big, but to be detail-oriented and concerned about methods. One of the skills I have developed during my PhD is scientific writing. Much of this development has come from writing collaboratively with Kay. For example, I have learned from Kay to never underestimate the value of telling a clear and coherent story. I am also very thankful to Kay for numerous exciting opportunities, such as building connections with local, national, and international collaborators, and finding funding for my research and professional travel. One such example was when Kay and I developed the Collaborative Outcomes Study of Meta-analyses in HIV/AIDS (COSMAHA), based at the South African Cochrane Centre in Cape Town, South Africa. Funded through the Johns Hopkins Center for Global Health and Kay's

discretionary funds, I was able to spend two months in Cape Town, collaborating with members of the South African Cochrane Centre and the Cochrane Review Group on HIV/AIDS (both based in Cape Town). This project, besides providing the data for Aim 1 of my dissertation, helped me expand my international network of collaborators, something that will definitely be invaluable to me in the future.

Before I started my PhD in 2011, I worked for about five years with Dr. Karen Robinson at the Johns Hopkins School of Medicine. Karen is a true inspiration and a professional par excellance. It is from working with her and others at the Johns Hopkins Evidence-based Practice Center (EPC) that I fell in love with systematic reviews and the methods used to conduct them. She was very supportive of me applying for a PhD and has been a wonderful mentor, always encouraging me to aim higher!

I am very thankful for the members of my PhD dissertation committee: Drs. Kay Dickersin, Roberta Scherer, and Marie Diener-West. They have been incredibly supportive. I also thank all members and alternate members of my departmental and school-wide preliminary oral examination committees (Drs. Marie Diener-West (Chair), Kay Dickersin, Roberta Scherer, Stephan Ehrhardt, Henry Jampel, Karen Robinson, Janet Holbrook, Joanne Katz, and Lawrence Moulton) and all thesis readers and members and alternate members of my final oral examination committee (Drs. Joanne Katz (Chair), Kay Dickersin, Roberta Scherer, Henry Jampel, Marie Diener-West, and Stephan Ehrhardt). The questions

that I was asked and the suggestions I received during these examinations were incredibly thoughtful and useful.

One of the unwritten prerequisites of PhD programs is a love for learning. During my program, I have enjoyed learning from a number of phenomenal teachers. Teachers at the Bloomberg School of Public Health do not just teach; they inspire. I am indebted to each one of them for pushing me to greater things, both during courses that I took and those that I assisted with teaching. I am also very thankful for everything that I have learned from various faculty members, staff, and fellow students in the Department of Epidemiology, the Center for Clinical Trials and Evidence Synthesis, and the Cochrane Eyes and Vision Group US Satellite – Dr. Kay Dickersin, Dr. Tianjing Li, Dr. Roberta Scherer, Dr. Barbara Hawkins, Dr. Swaroop Vedula, Dr. Evan Mayo-Wilson, Dr. Susan Hutfless, Dr. Asieh Golozar, Mrs. Kristina Lindsley, Dr. Shilpa Viswanathan, Mrs. Nicole Fusco, Ms. Gillian Gresham, Mr. Jimmy Le, and many others. At numerous research-in-progress meetings and journal clubs, they have listened patiently and provided important constructive criticism as I bounced off ideas and presented my findings. They have always done so with a smile, and have taught me to love my work and everything around it to the fullest extent!

I am very grateful for Mrs. Frances Burman and others in the Academic Support Core in the Department of Epidemiology. Fran is an extraordinarily valuable resource to everybody in our department. She is very knowledgeable

about most administrative issues students need help with, and her door is always open.

Pursuing one's PhD at the Johns Hopkins Bloomberg School of Public Health is not cheap. I am very grateful for the funding support and research assistantship that I have received from the Cochrane Eyes and Vision Group US Satellite and the National Eye Institute, without which pursuing a four year program of this nature would have been almost impossible. I am grateful for the generosity of donors of the Department of Epidemiology awards – the Dorothy and Arthur Samet Student Support Fund award and the Miriam E. Brailey Fund award. I also am grateful for conference travel support from the Cochrane Eyes and Vision Group, the Center for Clinical Trials and Evidence Synthesis, the Center for Global Health, the Department of Epidemiology, and the Johns Hopkins Bloomberg School of Public Health Student Assembly. I am also thankful for the Thomas C. Chalmers Award for Best Oral Presentation at the 22nd Cochrane Colloquium.

As much as I have loved working on my PhD, we all need our outlets to unwind and re-charge. I am very thankful for my family and close friends, both near and far, for providing this to me. They are too numerous to name here, but each contributed greatly. I confide in them, draw my energy from them, and love them dearly. My friends in Baltimore and my brothers Carl and Terence deserve special mention. Being the youngest of three brothers who lost their parents relatively young, I have always felt loved and cared for by my brothers. Day or

night, they are always a phone call away, willing to lend an ear while I vent, and provide unflinching support and encouragement.

Last, and certainly not the least, I am eternally grateful for my late parents Lolita and Sylverius Saldanha, whom I loved to no end. While I could write a whole book about them and the inspiration they provided, not a word of this dissertation book would have been written if it were not for them. I am because they were. My mother Lolita deserves special mention because a few months after she retired as a professor at age 60, she began her own PhD program in Education at the University of Mumbai, something she had always regretted not doing. However, she never let retirement come in the way of an unfulfilled dream. Unfortunately, death came in the way two years later, even before her preliminary oral examination.

To my mother Lolita and my father Sylverius: this one's for you!

## TABLE OF CONTENTS

Page

## LIST OF ABBREVIATIONS

- AHRQ = Agency for Healthcare Research and Quality
- AMD = age-related macular degeneration
- ARVO = Association for Research in Vision and Ophthalmology
- CENTRAL = Cochrane Central Register of Controlled Trials
- CER = comparative effectiveness research
- CEVG = Cochrane Eyes and Vision Group
- CI = confidence interval
- COMET = Core Outcomes Measures for Effectiveness Trials
- CONSORT = Consolidated Standards of Reporting Trials
- CRD = Centre for Reviews and Dissemination
- CRG = Cochrane Review Group
- DR = diabetic retinopathy
- ICTRP = International Clinical Trials Registry Platform
- IOM = Institute of Medicine
- IQR = interquartile range
- iQWiG = Institute for Quality and Efficiency in Health
- IRB = Institutional Review Board
- LILACS = Latin American and Caribbean Health Sciences Literature
- nNBC = normalized node betweenness centrality
- OMERACT = Outcome Measures in Rheumatology
- PCORI = Patient-Centered Outcomes Research Institute
- PICO = population, intervention, comparison, outcomes
- PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-analyses
- PRISMA-P = Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocols
- PROMIS = Patient-Reported Outcomes Measurement Information System
- QALYs = quality-adjusted life years
- RCT = randomized controlled trial
- RR = relative risk
- SRDR = Systematic Review Data Repository

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER ONE

## Introduction

## *1. Background*

Outcomes matter. In clinical intervention research, outcomes are events or measures in study participants that are used to assess the effectiveness and/or safety of the intervention being studied.[1] In other words, whether or not an intervention is deemed to be effective and safe is chiefly determined by the outcomes examined and reported in the studies informing the intervention's use. *Randomized controlled trials* and *systematic reviews* of randomized controlled trials are the studies that provide the strongest form of evidence for or against the effectiveness of clinical interventions.[2] Outcomes make up a critical component of research questions in clinical trials and systematic reviews. In formulating research questions, researchers define the population, intervention, comparison, and outcomes (PICO) to be examined.

While the criticality of outcomes to research has been recognized for a while, in recent years increased attention has been paid to the fact that research should examine outcomes that patients consider important. The motivation behind this increased attention is that those who receive an intervention (patients) should know how the intervention might affect them in ways that they care about. With this in mind, the Patient-Centered Outcomes Research Institute (PCORI) was established in the United States as part of the Affordable Care Act of 2009 to fund *comparative effectiveness research* (CER). The Institute of Medicine defines CER as "generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or

to improve the delivery of care".[3] Much of PCORI's efforts have focused on ensuring that outcomes important to patients are examined and reported in research.[4]

## *2. Conceptual framework for outcome use in clinical intervention research*

The process of using the right outcomes in clinical intervention research has four main steps - outcome selection, outcome specification, data collection and analysis, and outcome reporting (see Figure 1-1). Each step has its own set of important considerations.

### *2.1 – Step 1: Selection of outcomes for a clinical intervention study*

The first step in using outcomes in a clinical intervention study is to select the right outcomes to examine. This step helps determine whether the intervention being evaluated is considered to be effective and safe.

#### *2.1.1 – Considerations during selection of outcomes for a clinical intervention study*

Clinical intervention researchers generally consider a number of factors in selecting outcomes for their studies. These factors include relevance of the outcomes to clinical practice; importance to patients; importance to policy-makers; measurability and implementation issues; and costs and feasibility (e.g., timing and number of visits required to capture changes in the outcome).[5-7]

Clinical relevance of outcomes is generally considered by those designing clinical trials and systematic reviews. For many conditions, a patient's response to a clinical intervention could manifest in a variety of ways. To capture the various effects of a clinical intervention on symptoms and signs, physiological measurements, biomarkers, and adverse events, researchers may examine numerous outcomes among participants.[8] It is important that the outcomes examined capture both how effective (*efficacy outcomes*) and how safe (*safety outcomes*) the intervention is. Knowledge of results pertinent to both types of outcomes is critical in determining whether the possible benefits of the intervention outweigh the risks.

Clinical intervention researchers and those who fund research often consider certain outcomes useful because they can shorten the duration of a trial or follow-up. These outcomes are known as *interim outcomes* because they are assumed or hypothesized to be on the causal pathway to a more clinically-relevant outcome (e.g., CD4 counts in HIV-positive patients). Because it is often not known for certain whether interim outcomes are on a causal pathway or are simply risk factors (e.g., high intraocular pressure as a risk factor for glaucoma), these interim outcomes are often termed *surrogate outcomes*.

Given the importance of the patient perspective, clinical intervention researchers are increasingly being encouraged to include more outcomes important to patients. *Patient-important outcomes*, also known as *patient-centered outcomes*,[4] are outcomes that patients value directly (e.g., relief from

symptoms), which often contrast with physiological outcomes that clinicians might consider relevant but patients feel are of less immediate concern or not meaningful (e.g., intraocular pressure).[9] Other examples of patient-important outcomes include confidence in driving at night or fatigue. In addition, effectiveness and safety might be of different importance among patients (for example, depending on a patient's age and perceived severity of the disease), or between clinicians and patients.

The preferences of policymakers and healthcare payers also may be considered when selecting outcomes. They are generally interested in an intervention's cost and cost-effectiveness (e.g., quality-adjusted life-years (QALYs)).

*2.1.2 – Value of core outcome sets*

Depending on the various considerations described above, different researchers might reasonably arrive at different sets of outcomes. Such variation in outcomes across studies evaluating a clinical intervention's effectiveness and safety makes it challenging to determine what intervention works best for a certain condition.[10] For example, hundreds of measurement scales have been used to assess mental status [11] and quality-of-life,[12] making meaningful comparisons across studies challenging. Such variation also threatens credible evidence synthesis, both looking at a single intervention and across interventions for a

single condition, affecting clinicians, patients, clinical practice guideline developers, and systematic reviewers.

Efforts to promote comparability of outcomes across related studies in a field have led to the creation of *core outcome sets*.[13-15] One such effort is the Core Outcomes Measures for Effectiveness Trials (COMET) Initiative. COMET, launched in 2010, endeavors to develop methods for and collate core outcome sets. COMET defines core outcome sets as the minimum outcomes that should be examined in all clinical trials addressing a given condition.[16]

*2.1.3 – Limitations of current approaches to developing core outcome sets*

While core outcome sets could be immensely valuable, current approaches to develop them vary. A recent systematic review found a variety of disparate approaches to developing core outcome sets, including semi-structured group discussions (e.g., workshops), unstructured group discussions, literature reviews, and surveys.[17] The authors of that systematic review report on the lack of a current best practice, and suggest that the credibility of a core outcome set depends on the use of sound methods. Because systematic methods to identify core outcome sets are variable, it stands to reason that the outcomes selected are inconsistent across clinical trials and systematic reviews. In a recent study in the field of pre-term births, more than one-third of the 174 systematic reviews and more than one-third of the 1041 clinical trials studied did not examine chronic lung disease, a major clinical outcome relevant to that patient population.[18]

The clinical intervention research community is interested in identifying the "best" approach to developing core outcome sets.[17] In Aim 1 of this dissertation, we examine whether a new application of an existing method from the social sciences, *social network analysis*, can be used to identify *central* outcomes, which, in turn, could be used to develop core outcome sets (see objective and specific aims of this dissertation, page 21). Central outcomes refer to outcomes that are important to connecting other outcomes in a network of outcomes.

### *2.1.4 – Social network analysis as a method to identify central outcomes for core outcome sets*

Social network analysis is the study of graphs as representation of relationships and patterns of interaction among nodes within a network.[19] The methods for social network analysis, with their basis in network theory,[20] have increasingly been applied to health-related research. Some of these applications include evaluating collaboration among researchers in cardiovascular cohort studies;[21] evaluating scholarly citation patterns in Alzheimer's disease;[22] and evaluating the associations between personal relationships (e.g., spouses, friends) and happiness,[23] depression,[24] food choices,[25] physical activity,[26] alcohol consumption,[27] marijuana use,[28] and smoking.[29-31]

Social network analysis provides methodological tools that could be used to understand patterns of co-occurrence of outcomes in a given field. The various outcomes examined within a group of related clinical trials or systematic reviews

constitute a *network* of outcomes. When a single study examines more than one outcome, outcomes (or *nodes*) are said to *co-occur*. All these outcomes might be examined within a network of clinical trials or systematic reviews, but not necessarily within a single clinical trial or systematic review. Understanding the affinity (or repulsion) between certain outcomes that results in their co-occurrence (or not) would help identify outcomes that are considered *central* (i.e., more important relative to other outcomes in a network) to a network of outcomes in clinical trials or systematic reviews. These central outcomes could then be used, in tandem with frequent outcomes, as a starting point for developing core outcome sets. To our knowledge, social network analysis methods have not been used to identify central outcomes for core outcome sets.

### *2.2 – Step 2: Specification of outcomes for a clinical intervention study*

#### *2.2.1 – Importance of complete pre-specification of outcomes in clinical intervention research*

Once an outcome is selected for use in a clinical intervention study, the next step is completely specifying the outcome before the study is started. Pre-specification of outcomes is important for a number of reasons. First, it promotes transparency and reproducibility of the research process. Second, when outcomes are pre-specified, others can learn whether the authors have selectively reported the study outcomes, which can result in bias if reporting is based on the direction of the results (*outcome reporting bias*).[32-35] Third, pre-specifying outcomes

helps researchers plan the study, measure and analyze the data, and report results for all outcomes that they consider important, thereby minimizing waste of resources on outcomes that they will not use. Indeed, pre-specification of outcomes at the protocol stage has been recommended for both clinical trials [36, 37] and systematic reviews.[38-40]

*2.2.2 – How should outcomes be specified in clinical intervention research?*

Complete specification of an outcome (for example, in a study protocol) involves more than just specifying *what* outcomes will be examined during a clinical intervention study. It is also important to specify *how* and *when* the outcome will be examined during the study. ClinicalTrials.gov, an online register of ongoing clinical trials, recommends that clinical trialists specify each outcome using four elements - (1) the *domain* or outcome title; (2) the *specific measurement* or technique/instrument used to make the measurement; (3) the *specific metric* or format of the outcome data from each participant that will be used for analysis; and (4) the *method of aggregation* or how data from each group will be summarized.[37] In addition, the *time-points* that will be used for analysis must be specified.[37] For example, consider the common ophthalmologic outcome visual acuity (*domain*). The clinical trialist might choose to use the Snellen chart (*specific measurement*) to measure visual acuity as a change from baseline (*specific metric*). Finally, the clinical trialist might be interested in combining visual acuity data from individual participants in each arm of the trial

using the mean (*method of aggregation*) at 6 months (*time-point*). Clearly, it is possible that the study results presented for the outcome visual acuity might be different if using a 12 month time-point of outcome measurement instead of 6 months, or if the value of mean Snellen visual acuity at one of those time-points is used as specific metric instead of change in mean Snellen visual acuity from baseline.

*2.2.3 – Current state of outcome specification in clinical intervention research*

Despite existing recommendations on how outcomes should be examined and reported in clinical intervention research,[36, 38, 40, 41] complete pre-specification of outcomes in clinical trials is unsatisfactory. Zarin et al. reported that among 100 randomly-selected clinical trials registered on ClinicalTrials.gov, only the domain was pre-specified for 36% of outcomes; only the domain and specific measurement were specified for 25%; and only the domain, specific measurement, and specific metric were specified for 26%.[37]

We also know that it is currently unclear how systematic reviewers select outcomes, and to what extent the outcomes they select are pre-specified. While formulating the research question for a systematic review, the systematic reviewer defines the PICO to be examined. Studies eligible for a systematic review address the review question and are broadly similar with regard to the population, intervention, and comparison groups. However, the studies included in the systematic review frequently report different outcomes from one another and those

chosen by the systematic reviewer.[18] If authors of systematic reviews change which outcomes are designated as primary or secondary based on the results in the included studies, this can lead to biased reporting of findings in systematic reviews.[42]

The research community must agree on three important related issues - (1) how outcomes should be selected for a systematic review (e.g., should they reflect the outcomes in clinical trials or should the systematic reviewer select outcomes *de novo*?); (2) a core outcome set that should be examined and reported in all clinical trials and systematic reviews in a field (assuming this set is updated as needed); and (3) the elements necessary for complete pre-specification and reporting of outcomes.

In Aim 2 of this dissertation, we study a sample of Cochrane systematic reviews addressing four major eye conditions (glaucoma, cataract, age-related macular degeneration (AMD), and diabetic retinopathy (DR)). We study these systematic reviews to examine aspects of items (2) and (3) above - completeness of outcome pre-specification and comparability of outcomes (see objective and specific aims of this dissertation, page 21).

### *2.3 – Step 3: Data collection and analysis of outcomes in a clinical intervention study*

After pre-specifying the outcomes to be examined, clinical researchers implement their plans, collecting and analyzing the data. Important considerations

include minimizing measurement error (information bias), missing data, and bias in the analysis. These considerations, while vital to the validity and reliability of a clinical intervention study, are not the direct focus of this dissertation. However, complete pre-specification (as discussed in step 2), helps minimize the potential for missing data and biased analysis.

### *2.4 – Step 4: Reporting of outcomes from a clinical intervention study*

The two main considerations when reporting findings from a clinical intervention study are reporting dependable results and avoiding reporting bias.

#### *2.4.1 – Reporting dependable results*

Researchers should strive to report results transparently, accurately, and dependably, whatever the forum. Dependability of reported results is compromised when discrepancies exist in results for the same outcome in various documents (e.g., protocol, trial registry entries, full publication) of the same clinical trial. Such discrepancies call the validity of the clinical trial and the veracity of its reports into question.

Reported results of clinical trials are generally available as published journal articles and/or from unpublished sources (i.e., *grey literature*). Among unpublished sources of clinical trial results, conference abstracts are the most important.[43] For one, although there have been recent advances related to open access to findings of scientific research,[44] only about 20% of the scientific

literature is available freely as open access material.[45] This increases the value of information presented at conferences, especially for individuals, such as community-based clinicians or those in resource-poor settings, who might not have access to publications requiring paid access (e.g., a journal subscription).

Research has documented discrepancies in reported results between conference abstracts and corresponding full publications. For as many as 40%-60% of clinical trials in the fields of orthopedics,[46, 47] cardiology,[48] pediatrics,[49] pediatric surgery,[50] and infectious disease,[51] discrepancies exist comparing results reported in conference abstracts and full publications.

That discrepancies exist to this extent is troubling for two main reasons. First, clinical decisions are sometimes based solely on results presented as conference abstracts.[50, 52, 53] For example, Gross and colleagues demonstrated that pre-publication dissemination of results promptly led to substantial changes in clinical practice associated with carotid endarterectomy.[52] Second, it is recommended that systematic reviewers include results reported only in conference abstracts.[3, 38, 40] If what is reported in conference abstracts is not accurate, then inclusion of the data is worrisome.

In the first part of Aim 3 of this dissertation, we examine the amount of agreement in main outcome results between conference abstracts and corresponding full publications for the same clinical trial (see objective and specific aims of this dissertation, page 21). To our knowledge, this has not been

previously examined using a large sample of randomized controlled trials in ophthalmology.

*2.4.2 – Avoiding reporting bias*

Ideally, all outcomes measured in a clinical intervention study are analyzed ("results") and reported by the study investigators. In clinical trials, under-reporting of results has been called "scientific misconduct".[54-57] Because clinical trial participants volunteer to participate in clinical trials with the understanding that their participation advances science, under-reporting is a violation of that understanding.[55, 57-59] Indeed, this is embodied in the general principles of the Declaration of Helsinki which states, "Researchers have a duty to make publicly available the results of their research on human subjects and are accountable for the completeness and accuracy of their reports".[59]

Under-reporting of results from clinical trials does not lead to bias *per se* if the results under-reported are a random subset of all existing results. However, when the nature and direction of the results influences their reporting, this is called *reporting bias* [38, 58] or, more broadly, *dissemination bias*.[35] As part of this dissertation, we focus on two distinct but related reporting biases - *publication bias* and *outcome reporting bias*.

*2.4.2.1 – Publication bias*

*Publication bias* refers to the tendency among researchers, peer reviewers, and journal editors to submit or accept manuscripts for publication based on the direction or strength of the study results.[35, 58, 60, 61]

Conference abstracts are subject to publication bias, which is worrying if they are used as a source of evidence by decision makers and systematic reviewers (e.g., when these individuals do not have access to a full publication or when study findings have not been published in full). Conference abstracts describing statistically significant treatment effects or those with positive results are more likely to be published in full.[35, 57, 62] In a Cochrane systematic review, Scherer et al. reported that conference abstracts of clinical trials with positive results were 18% more likely to be published in full.[57] Even when there is a full publication, presentation at a conference has been shown to be associated with delays in publication,[63] and clinical trials with positive results have been shown to be published in full sooner than those without positive results (i.e., *time lag bias)*.[64-66] Trial quality and sample size have also been shown to be associated with full publication.[57]

Research has also shown that among clinical trials presented at conferences, clinical trials funded by industry are more likely to reach full publication than those not funded by industry.[57] The small number of conference abstracts reporting funding source, statistical significance of the findings (in terms of the association between the intervention and primary outcome examined), and

information about whether they reached full publication made it impossible to examine how positive findings are related to the apparent influence of funding on full publication. Among published clinical trials, clinical trials funded by industry are more likely than non-industry funded clinical trials to report results favoring the funder.[67-70] It would be interesting to see whether this relationship is also present in the case of full publication of conference abstracts.

The impact of study investigators' *conflicts of interest* on publication is even less clear. Conflicts of interest refer to the set of conditions in which professional judgment concerning a primary interest (such as a patient's welfare or the validity of research) might be unduly influenced by a secondary interest (such as financial gain).[71, 72]

In published clinical trials, the presence of financial conflicts of interest has been shown to be associated with greater likelihood of the authors' conclusions favoring experimental interventions.[73] However, the impact of author conflicts of interest on study publication itself appears complex. On one hand, financial gain might threaten impartial judgment,[74] leading to study authors selectively publishing studies (*publication bias* [35,60, 61]), delaying publication of studies (*time lag bias* [64-66]), or selectively reporting certain outcomes (*outcome reporting bias* [32-35, 75]), based on the nature and direction of the results. On the other hand, having a financial relationship with industry, and the consequent incentives, might work to facilitate publication, irrespective of the direction of study results. Currently, a knowledge gap exists as to what the association is

between clinical trial author conflicts of interest and full publication of results, and whether direction of results has any bearing on this association. In the second part of Aim 3, we address this gap by examining the association between author conflicts of interest and full publication of results of clinical trials presented as conference abstracts (see objective and specific aims of this dissertation, page 21).

*2.4.2.2 – Outcome reporting bias*

Outcome reporting bias typically is manifested by an overestimation of treatment effects.[32-35, 75-78]

There are three main scenarios under which outcome reporting bias can occur. Each of these scenarios may be due to deliberate attempts to obfuscate or due to a genuine belief that the action is acceptable, or reasons between these two extremes. First, outcome reporting bias can occur because of *selective outcome reporting*, i.e., selective reporting on the basis of the results of a subset of the original pre-specified outcomes.[34, 78] Second, outcome reporting bias can occur when what is reported to be the primary outcome in the publication is different from the pre-specified primary outcome. This might be done to provide emphasis to a certain statistically significant or favorable non-primary outcome by making it the primary outcome.[79] Research has shown that for between 33% and 67% of clinical trials, the primary outcome in the protocol differed from that in the final publication,[33, 34, 75, 78, 80] and that for 31% of clinical trials the primary outcome changed between trial registration and publication.[79] Changing of

primary outcomes, addition of new outcomes in the full publication, and deletion of outcomes presented in the conference abstract may all occur as a consequence of statistical significance of results of specific outcomes.[57] Third, outcome reporting bias can occur even if all outcomes are reported, but the level of detail reported for outcomes depends on the statistical significance of the results. For example, in a sample of 101 clinical trials evaluated by the Institute for Quality and Efficiency in Health (iQWiG), publicly available sources had complete information on only 39% of outcomes.[81] In a random sample of 283 Cochrane systematic reviews, for more than half the reviews (55%), the review authors could not include full data for the review primary outcome of interest from all eligible clinical trials.[82] Studies that have compared the results of meta-analyses based on published data with individual patient data meta-analyses addressing the same research question have found that, compared with using individual patient data, using only published data provides up to three times larger effect estimates and, in some instances, leads to qualitatively different conclusions.[83-85]

Taken together, both parts of Aim 3 address two important related issues. In the first part, we examine whether conference abstracts are dependable, especially in scenarios where a conference abstract is not published in full or if the full publication is not found. In the second part, we address a specific publication bias issue - whether author conflicts of interest are associated with a conference abstract reaching full publication.

### *2.4.3 – Advantages and disadvantages of including data from conference abstracts in systematic reviews*

Systematic reviewers strive to be comprehensive in their endeavor. However, the validity of the results obtained and the strength of the conclusions made in systematic reviews are affected by the potential for reporting biases and the dependability of reported data in the included studies. Including results from conference abstracts promotes comprehensiveness in the systematic review process, but systematic reviewers should recognize the advantages and disadvantages of including data from conference abstracts.

Including conference abstracts has some important advantages. Conference abstracts can represent the most recent research results, thus providing systematic reviewers access to clinical trial results that might not yet be available elsewhere. In some instances, conference abstracts might be the only ever available source of information about a clinical trial (e.g., if the results are never published in full or are published in full in a language or database that is not accessible to the systematic review authors).

Including conference abstracts in systematic reviews has its disadvantages, however. First, abstracts submitted for presentation at conferences typically do not undergo peer review. Peer review refers to the use of experts, or author peers, to help judge the value of submitted work.[86] These experts or peers critically evaluate the creative work by others in the same field. Second, conference abstracts often contain preliminary results and do not include final analyses.

Results from preliminary and final analyses might differ, especially if the preliminary analysis is based on a subset of the entire study population or if additional analyses are subsequently conducted. Third, systematic reviewers need to guard against "double counting" results of clinical trials presented both as conference abstracts and full publications. This could occur if systematic reviewers fail to identify a conference abstract and a full publication as addressing the same clinical trial, either because the reported information (e.g., clinical trial registration number, authors, sample size, intervention details) differs in the two documents, or because there is inadequate detail in either document to conclusively identify the match. This double counting could affect the validity of estimated meta-analytic effect estimates and/or falsely increase their precision. Fourth, authors of conference abstracts are usually constrained by the amount of space (number of words) available. Adherence to this constraint can come at the expense of an explicit and detailed description of methods used to conduct the clinical trial. This limits the ability of the reader of a conference abstract to critically assess the methods, evaluate risk of bias, and conduct a qualitative synthesis of the clinical trial in the context of the entire body of evidence.[87] Fifth, researchers and representatives of pharmaceutical companies may view conferences as opportunities to advertise recent results of clinical trials of drugs and devices to clinicians and members of professional clinical organizations. This can be problematic because these results might not yet or might not ever be peer reviewed or published, and even if they are published, research has shown that

pharmaceutical companies often selectively report outcome analyses and results, leading to outcome reporting bias.[67, 88] Sixth, searching for and obtaining conference abstracts is time-consuming, especially if the conference abstracts are not indexed by electronic databases such as MEDLINE.[89]

### *3. Objective and specific aims of this dissertation*

The overarching objective of this dissertation is to address important questions and research gaps related to selecting, specifying, measuring, analyzing, and reporting outcomes in clinical intervention research. To achieve this, three specific aims are listed below and described in detail in the following three chapters of this dissertation:

- *Aim 1 - Use social network analysis methods to (a) understand patterns of co-occurrence of outcomes in systematic reviews of HIV/AIDS; and (b) identify outcomes that are central to the network of outcomes examined in systematic reviews of HIV/AIDS.*
- *Aim 2 - Assess the completeness of pre-specification and comparability of outcomes in systematic reviews addressing four common eye conditions.*
- *Aim 3 – Evaluate, using randomized controlled trials in ophthalmology, (a) the agreement in reported main outcome results comparing abstracts and their corresponding full publications; and (b) the association between author conflicts of interest and full publication of results presented in abstracts.*

### *4. Datasets used in this dissertation*

For Aim 1, we have elected to examine systematic reviews of interventions to prevent and treat HIV/AIDS because of the variety of interventions and outcomes inherent to the field. This variety lends itself nicely to the understanding of patterns of co-occurrence of outcomes. We include all 140 Cochrane systematic reviews and systematic review protocols published by the Cochrane Review Group on HIV/AIDS in the *Cochrane Database of Systematic Reviews* as of June 30, 2013 (Issue 6). For completed systematic reviews (n=99), we include the most recent version. For systematic reviews in progress (n=41), we include the protocol (a peer reviewed and published document outlining the planned methods [including outcomes to be examined]).

For Aim 2, we select four common eye conditions in ophthalmology (glaucoma, cataract, age-related macular degeneration (AMD), and diabetic retinopathy (DR)), because of their high disease burden across populations.[90] We include all 57 Cochrane systematic reviews protocols that addressed these four conditions and were published by the Cochrane Eyes and Vision Group (CEVG) in *The Cochrane Database of Systematic Reviews* as of June 30, 2013 (Issue 6). For systematic reviews for which we were able to find protocol (n=54), we include the protocol. For the remaining systematic reviews (n=3), we include the Methods section of the completed versions of the systematic review.

For Aim 3, we include all 513 conference abstracts describing results of randomized controlled trials presented at the 2001-2004 Association for Research

in Vision and Ophthalmology (ARVO) conferences. This annual conference is the largest international research conference in ophthalmology in the world. Eligible for this study are randomized controlled trials addressing any type of intervention for any eye condition or in healthy volunteers.

### *5. Significance and innovation*

All clinical intervention research involves using outcomes to examine whether the intervention is effective and safe. This dissertation takes a holistic approach to addressing important issues related to outcome use in clinical intervention research. We recognize that the process of using outcomes begins well before researchers report the results, and that it includes the steps of selection, specification, data collection and analysis, and reporting of the outcomes (Figure 1-1). This dissertation addresses research gaps related to three of these steps.

To examine outcome selection, we conduct a case study examining the innovative application of social network analysis methods (Aim 1). The product of this application is a list of *central* outcomes, which, used in tandem with outcomes that are the most frequent, can help inform the development of core outcome sets.

Regarding the step of specifying outcomes, we adapt the outcomes specification framework used by ClinicalTrials.gov and propose that all clinical researchers, including systematic reviewers, use a five-element framework to specify each outcome (Aim 2). Currently, while all the five elements we propose are discussed in various systematic review methods guideline documents,[38-40]

individual elements are listed in various disparate locations in those documents. The idea that a five-element framework should be used for pre-specifying outcomes appears new to the systematic review community.

Finally, for reporting outcomes, we examine all conference abstracts describing randomized controlled trials presented at the largest research conference in ophthalmology in the world (ARVO) and their corresponding full publications (Aim 3). We evaluate the agreement in main outcome results between conference abstracts and full publications of the same randomized controlled trial, an issue important to clinicians, patients, and systematic reviewers. Also, we capitalize on the fact that ARVO collects conflicts of interest information from all authors who submit abstracts, and use that information to understand the association between conflicts of interest and full publication of randomized controlled trials.

## REFERENCES


1. Meinert CL. *Clinical trials dictionary: Terminology and usage recommendations*. 2nd edition. 2012. Wiley. Hoboken, NJ.

2. Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ, Green L, Naylor CD, Wilson MC, Richardson WS. Users' Guides to the Medical Literature: XXV. Evidence-based medicine: principles for applying the Users' Guides to patient care. Evidence-Based Medicine Working Group. *JAMA* 2000; 284(10):1290-6.

3. Institute of Medicine. Initial National Priorities for Comparative Effectiveness Research. 2009. Available online at: http://www.iom.edu/~/media/Files/Report%20Files/2009/ComparativeEffectivenessResearchPriorities/CER%20report%20brief%2008-13-09.pdf. Accessed July 17, 2015.

4. Patient-Centered Outcomes Research Institute. Patient-Centered Outcomes Research. Available online at http://www.pcori.org/research-we-support/pcor/. Accessed July 17, 2015.

5. DeMets D, Friedman L, Furberg CD. Counting events in clinical trials. *N Engl J Med*. 1980; 302(16):924-5.

6. Sinha IP, Altman DG, Beresford MW, Boers M, Clarke M, Craig J, Alberighi OD, Fernandes RM, Hartling L, Johnston BC, Lux A, Plint A, Tugwell P, Turner M, van der Lee JH, Offringa M, Williamson PR, Smyth RL; StaR Child Health Group. Standard 5: selection, measurement, and reporting of outcomes in clinical trials in children. *Pediatr* 2012 Jun; 129 Suppl 3:S146-52.

7. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012; 13(132).

8. Pocock SJ. Clinical trials with multiple outcomes: a statistical perspective on their design, analysis, and interpretation. *Control Clin Trials* 1997; 18(6):530-45.

9. JAMA Evidence. Glossary: Patient-important outcomes. Available online at: http://jamaevidence.mhmedical.com/glossary.aspx. Accessed July 17, 2015.

10. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007; 8:39.

11. Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ* 1998; 317(7167):1181–1184.

12. Salek S. Compendium of Quality of Life Instruments. 1999. Wiley. ISBN: 0-471-98145-1.

13. Tugwell P, Boers M, Brooks P, Simon LS, Strand V. OMERACT: An international initiative to improve outcome measures in rheumatology. *Trials* 2007; 8:38.

14. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, et al. Core outcome measures for chronic pain trials: IMMPACT recommendations. *Pain* 2005; 113(1-2):9-19.

15. Patient Reported Outcomes Measurement Information System. About PROMIS. Available online at http://www.nihpromis.org/about/abouthome. Accessed July 17, 2015.

16. COMET Initiative. COMET Initiative. Available online at: www.comet-initiative.org. Accessed July 17, 2015.

17. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, Williamson PR. Choosing health outcomes for comparative effectiveness research: A systematic review. *PLOS One* 2014. DOI: 10.1371/journal.pone.0099111.

18. Ioannidis JP, Horbar JD, Ovelman CM, Brosseau Y, Thorlund K, Buus-Frank ME, Mills EJ, Soll RF. Completeness of main outcomes across randomized trials in entire discipline: survey of chronic lung disease outcomes in preterm infants. *BMJ* 2015;350:h72. doi: 10.1136/bmj.h72.

19. Wasserman, S. & Faust, K. *Social Network Analysis Methods and Applications*. 1994. Cambridge: Cambridge University Press.

20. Freeman, L.C., 1979. Centrality in networks: I. Conceptual clarification. *Social Networks* 1, 215–39.

21. Eblen M, Fabsitz RR, Olson JL, Pearson K, Pool LR, Puggal M, et al. Social network analysis comparing researcher collaborations in two cardiovascular cohort studies. *Research Evaluation* 2012; 21: 392-405.

22. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. *BMJ* 2009; 339:b2680. doi: 10.1136/bmj.b2680.

23. Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ*. 2008;337:a2338. doi: 10.1136/bmj.a2338.

24. Rosenquist JN, Fowler JH, Christakis NA. Social network determinants of depression. *Mol Psychiatry* 2011; 16(3): 273-81.

25. Pachucki MA, Jacques PF, Christakis NA. Social network concordance in food choice among spouses, friends, and siblings. *Am J Public Health* 2011; 101(11): 2170-7.

26. de la Haye K, Robins G, Mohr P, Wilson C. How physical activity shapes, and is shaped by, adolescent friendships. *Soc Sci Med* 2011; 73(5): 719-28.

27. Rosenquist JN, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. *Ann Intern Med* 2010; 152(7): 426-33.

28. Tucker JS, de la Haye K, Kennedy DP, Green HD, Pollard MS. Peer influence on marijuana use in different types of friendships. *J Adolesc Health* 2014; 54(1): 67-73.

29. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *N Engl J Med* 2008; 358(21): 2249-58.

30. Valente TW, Unger JB, Johnson CA. Do popular students smoke? The association between popularity and smoking among middle school students. *J Adolesc Health* 2005; 27(4): 323-9.

31. Green HD Jr, Horta M, de la Haye K, Tucker JS, Kennedy DR, Pollard M. Peer influence and selection processes in adolescent smoking behavior: a comparative study. *Nicotine Tob Res* 2013; 15(2): 534-41.

32. Chan AW, Hrobjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004; 291(20):2457-2465.

33. Chan AW, Krleza-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ*. 2004;171(7):735-40.

34. Chan AW, Altman D. Identifying outcome reporting bias in randomized trials on PubMed: review of publications and survey of authors. *BMJ* 2005; 330(7494):753.

35. Song F, Parekh S, Hooper, L, Loke Y, et al. Dissemination and publication of research findings: an update of related biases. *Health Technol Assess* 2010;14(8).

36. ClinicalTrials.Gov. How to submit your results. Available online at: https://clinicaltrials.gov/ct2/manage-recs/how-report#StepsForSubmittingResults. Accessed July 17, 2015.

37. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database–update and key issues. *N Engl J Med* 2011; 364(9):852-860.

38. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. Available online at: www.cochrane-handbook.org. Accessed July 17, 2015.

39. Chandler J, Churchill R, Higgins J, Tovey D. *Methodological standards for the conduct of new Cochrane Intervention Reviews*. Version 2.2. 17 December 2012. Available: http://www.editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/MECIR_conduct_standards%202.2%201712 2012.pdf. Accessed July 17, 2015.

40. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov. Accessed July 17, 2015.

41. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010; 340:c869. doi: 10.1136/bmj.c869.

42. Kirkham JJ, Altman DG, Williamson PR. Bias Due to Changes in Specified Outcomes during the Systematic Review Process. *PLoS ONE* 2010; 5(3):e9810. doi: 10.1371/journal.pone.0009810.

43. Balshem H, Stevens A, Ansari M, Norris S, Kansagara D, Shamliyan T, Chou R, Chung M, Moher D, Dickersin K. Finding Grey Literature Evidence and *Assessing for Outcome and Analysis Reporting Biases When Comparing Medical Interventions: AHRQ and the Effective Health Care Program. Methods Guide for Comparative Effectiveness Reviews*. (Prepared by the Oregon Health and Science University and the University of Ottawa Evidence-based Practice Centers under Contract Nos. 290-2007-10057-I and 290-2007-10059-I.) AHRQ Publication No. 13(14)-EHC096-EF. Rockville, MD: Agency for Healthcare Research and Quality. November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm. Accessed July 17, 2015.

44. Liesegang TJ. The continued movement for open access to peer-reviewed literature. *Am Journal of Ophthalmol* 2013; 156(3):423–432.

45. Björk BC, Welling P, Laakso M, Majlender P, Hedlund T, Gudnason G. Open access to the scientific literature: situation 2009. *PLoS One.* 2010; 5(6):e11273.

46. Bhandari M, Devereaux PJ, Guyatt GH, Cook DJ, Swiontkowski MF, Sprague S, Schemitsch EH. An observational study of orthopaedic abstracts and subsequent full-text publications. *J Bone Joint Surg Am* 2002; 84-A:615–621.

47. Kleweno CP, Bryant WK, Jacir AM, Levine WN, Ahmad CS. Discrepancies and rates of publication in orthopedic sports medicine abstracts. *Am J Sports Med* 2008; 36(10):1875-9.

48. Toma M, McAlister FA, Bialy L, Adams D, Vandermeer B, Armstrong PW. Transition from meeting abstract to full-length journal article for randomized controlled trials. *JAMA* 2006; 295:1281–1287.

49. Klassen TP, Wiebe N, Russell K, Stevens K, Hartling L, Craig WR, Moher D. Abstracts of randomized controlled trials presented at the society for pediatric research meeting: an example of publication bias. *Arch Pediatr Adolesc Med* 2002; 156:474–479.

50. Weintraub WH. Are published manuscripts representative of the surgical meeting abstracts? An objective appraisal. *J Pediatr Surg* 1987; 22:11–13.

51. Rosmarakis ES, Soteriades ES, Vergidis PI, Kasiakou SK, Falagas ME. From conference abstract to full paper: differences between data presented in conferences and journals. *FASEB J* 2005; 19:673–680.

52. Gross CP, Steiner CA, Bass EB, Powe NR. Relation between prepublication release of clinical trial results and the practice of carotid endarterectomy. *JAMA* 2000; 284:2886–2893.

53. Falagas ME, Rosmarakis ES. Clinical decision-making based on findings presented in conference abstracts: is it safe for our patients? *Eur Heart J* 2006; 27(17):2038-9.

54. Chalmers I. Proposal to outlaw the term 'negative trial'. *BMJ* 1985;290:1002.

55. Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990; 263:1405–8.

56. Antes G, Chalmers I. Under-reporting of clinical trials is unethical. *Lancet* 2003; 361:978–9.

57. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev* 2007; Issue 2. Art. No.: MR000005. DOI: 10.1002/14651858.MR000005.pub3.

58. Dickersin K, Chalmers I. Recognizing, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon the WHO. *J R Soc Med* 2011;104(12):532-8.

59. World Medical Association. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA* 2013; doi: 10.1001/jama.2013.281053.

60. Simes RJ. Publication bias: The case for an international registry of clinical trials. *J Clin Oncol* 1986; 4:1529–41.

61. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990; 263: 1385-9.

62. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance of direction of trial result. *Cochrane Database Syst Rev* 2009; 1:MR000006.

63. Unalp A, Tonascia S, Meinert CL. Presentation in relation to publication of results from clinical trials. *Contemp Clin Trials* 2007; 28(4) 358-69.

64. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ* 1997; 315:640–645.

65. Dickersin K, Olson CM, Rennie D. Cook D, Flanagin A, Zhu Q, Reiling J, Pace B. Association between time interval to publication and statistical significance. *JAMA* 2002; 287(21):2829-31.

66. Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database Syst Rev* 2007; Issue 2. [DOI: 10.1002/14651858.MR000011.pub2]

67. Campbell EG, Louis KS, Blumenthal D. Looking a gift horse in the mouth: corporate gifts support life sciences research. *JAMA* 1998; 279:995-9.

68. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008; 358(3):252-60.

69. Sismondo S. Pharmaceutical company funding and its consequences: a qualitative systematic review. *Contemporay Clinical Trials* 2008; 29(2):109–13.

70. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2012; Issue 12. Art. No.: MR000033. DOI: 10.1002/14651858.MR000033.pub2.

71. Thompson D. Understanding financial conflicts of interest. *N Engl J Med.*1993; 329:573-576.

72. International Committee of Medical Journal Editors (ICMJE). Author responsibilities-Conflicts of interest. Available at http://www.icmje.org/recommendations/browse/roles-and-responsibilities/author-responsibilities--conflicts-of-interest.html. Accessed July 17, 2015.

73. Kjaergard LL, Gluud C. Citation bias of hepato-biliary randomized clinical trials. *J Clin Epidemiol* 2002; 55:407-10.

74. Johnston J. Conflict of Interest in Biomedical Research. In *From Birth to Death and Bench to Clinic: The Hastings Center Bioethics Briefing Book for Journalists, Policymakers, and Campaigns*, ed. Mary Crowley (Garrison, NY: The Hastings Center, 2008), 31-34.

75. Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *N Engl J Med* 2009; 361:1963-1971.

76. Furukawa TA, Watanabe N, Omori IM, Montori VM, Guyatt GH. Association between unreported outcomes and effect size estimates in Cochrane meta-analyses. *JAMA* 2007; 297:468e70.

77. Mathieu S, Boutron I, Moher D. Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009; 302(9):977-84.

78. Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, et al. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev* 2011*;* Issue 1. Art. No.: MR000031. DOI: 10.1002/14651858.MR000031.pub2.

79. Ewart R, Lausen H, Millian N. Undisclosed changes in outcomes in randomized controlled trials: An observational study. *Ann Fam Med* 2009; 7(6):542-6.

80. Dwan K, Gamble C, Williamson PR, Kirkham JJ, Reporting Bias Group. Systematic review of the empiral evidence of study publication bias and outcome reporting bias – an updated review. *PLoS One* 2013; 8(7):e66844. doi: 10.1371/journal.pone.0066844.

81. Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervolgyi V, et al. Completeness of reporting of patient-relevant clinical trial outcomes: comparison of unpublished clinical study reports with publicly available data. *PLOS Med* 2013; 10 (10) e1001526. doi: 10.1371/journal.pmed.1001526.

82. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010; 340:c365. doi: 10.1136/bmj.c365.

83. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993; 341:418-422.

84. McKormack K, Grant A, Scott N. Value of updating a systematic review in surgery using individual patient data. *Br J Surg* 2004; 91:495-9.

85. Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ* 2012; 344:d7202. doi: 10.1136/bmj.d7202.

86. Godlee F, Jefferson T. *Peer review in health sciences* (2nd ed.). BMJ Publishing Group. London, UK 2003.

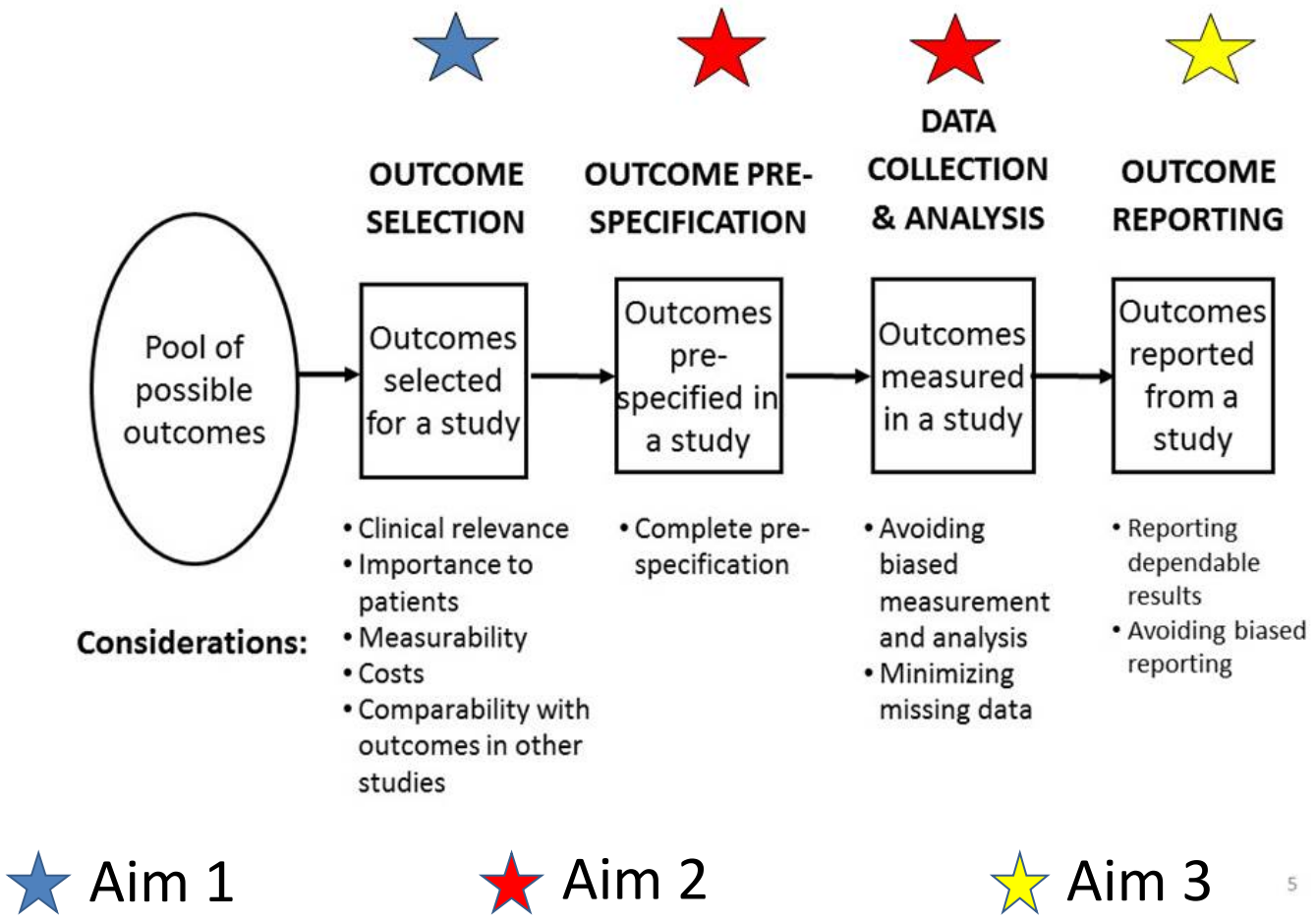
87. Dundar Y, Dodd S, Dickson R, Walley T, Haycox A, Williamson PR. Comparison of conference abstracts and presentations with full-text articles in the health technology assessments of rapidly evolving technologies. *Health Technol Assess* 2006; 10(5) :iii-iv, ix-145.

88. Vedula SS, Li T, Dickersin K. Differences in reporting of analyses in internal company documents versus published trial reports: comparisons in industry-sponsored trials in off-label uses of gabapentin. *PLoS Med* 2013; 10(1):e1001378. doi: 10.1371/journal.pmed.1001378. Epub 2013 Jan 29.

89. Dundar Y, Dodd S, Williamson P, Walley T, Dickson R. Searching for and use of conference abstracts in health technology assessments: policy and practice. *Int J Technol Assess Health Care* 2006; 22(3):283-7.

90. National Eye Institute. Statistics and Data. 2010. Available online at: http://www.nei.nih.gov/eyedata. Accessed July 17, 2015.

Figure 1-1: Conceptual framework demonstrating the four main steps of using outcomes in a clinical intervention study (clinical trial or systematic review)



# CHAPTER TWO

## MANUSCRIPT ONE

## Social network analysis can identify central outcomes potentially useful for core outcome sets: A case study of systematic reviews of HIV/AIDS

**Social network analysis can identify central outcomes potentially useful for core outcome sets: A case study of systematic reviews of HIV/AIDS**

Ian J Saldanha,[1] Tianjing Li,[1] Cui Yang,[2] Cesar Ugarte-Gil,[3] George W Rutherford,[4] Kay Dickersin[1]

[1] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

[2] Department of Health, Behavior, and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

[3] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

[4] Department of Epidemiology and Biostatistics, University of California, San Francisco

## ABSTRACT

### Introduction

Methods to develop core outcome sets (COS), the minimum outcomes that should be measured in research in a topic area, vary. We applied social network analysis methods to understand outcome co-occurrence in HIV/AIDS systematic reviews (SRs), and identify outcomes *central* to the network of outcomes in HIV/AIDS.

### Methods

We examined all Cochrane SRs of HIV/AIDS as of June 2013. We defined a *tie* as two outcomes (*nodes*) co-occurring in ≥2 SRs. To identify central outcomes, we used *normalized node betweenness centrality* (nNBC) (the extent to which connections between other outcomes in a network rely on that outcome as an intermediary). We conducted a subgroup analysis by HIV/AIDS intervention type (i.e., clinical management, biomedical prevention, behavioral prevention, and health services).

### Results

The 140 included SRs examined 1140 outcomes, 294 of which were unique. The most central outcome overall was all-cause mortality (nNBC=23.9). The most central and most frequent outcomes differed overall and within subgroups. For example, adverse events (specified), was among the most central but not among the most frequent outcomes, overall.

## Discussion

Social network analysis methods are a novel application to identify central outcomes, which provides additional information about influential outcomes, potentially useful for developing COS.

## INTRODUCTION

In clinical research, an outcome is an event or measure in study participants that is used to assess the effectiveness and/or safety of the intervention being studied.[1] What is typically thought of as the name of the outcome (e.g., "anxiety", "death"*),* is more formally called the outcome *domain*.[2, 3] In this paper, when we say "outcome", we are referring specifically to the outcome domain. Clinical trialists and systematic reviewers typically examine multiple outcomes in their studies, and the number of outcomes they examine varies widely. For example, clinical trials have been reported as examining between 1 and 71 primary outcomes and between 0 and 122 secondary outcomes.[2] This variation in outcomes examined creates inconsistent outcome reporting, and is a threat to credible evidence synthesis because when an outcome is reported for one trial but not another, it is impossible to compare or synthesize results across trials, for example, in a meta-analysis.

The various outcomes examined within a group of related clinical trials or systematic reviews constitute a *network* of outcomes. The network is made up of individual outcomes representing various categories (e.g., clinical/biological outcomes, behavioral outcomes). For example, a clinical trial assessing the efficacy of a statin for treatment of hypercholesterolemia might examine the clinical/biological outcomes serum cholesterol concentration and stroke. Some less frequently-examined, but not less relevant, outcomes might be cost-effectiveness, patient preferences, and quality-of-life. All these outcomes might be

examined within clinical trials or systematic reviews, but all would not typically be examined within a single clinical trial or systematic review. For example, a study addressing mother-to-child transmission of HIV might examine a cluster of outcomes such as the child's acquisition of HIV, premature delivery, neonatal morbidity, and neonatal mortality, outcomes that could also be examined in other studies on the topic. Clinical trialists and systematic reviewers alike would want information on relevant outcomes in that cluster to be present in all studies on the topic, yet this clustering may not be revealed by consensus, survey, literature review, or other methods of deciding on outcomes important to use across all studies on a topic.

When a single study examines more than one outcome, outcomes (or *nodes*) are said to *co-occur*. Within a given topic area, understanding the underlying co-occurrence of outcomes in existing research could inform the development of core outcome sets. A core outcome set refers to the minimum outcomes that should be examined in all clinical trials addressing a specific condition.[4] Core outcome sets are presumed to serve two main purposes: they (1) facilitate decision-making by patients, clinicians, healthcare payers, and guideline developers by promoting consistency in the outcomes examined in the research informing care, and (2) reduce the potential for selective reporting of an outcome purely on the basis of results.[3-7]

Social network analysis, the study of graphs as representation of relationships and patterns of interaction among *nodes* (in our case, outcomes)

within a network,[8] provides methodological tools to understand patterns of co-occurrence of outcomes in a given topic area. These patterns, along with other information, could contribute to development of core outcome sets. To our knowledge, social network analysis methods have not been applied in the context of analysis of co-occurrence of outcomes or to identify potential outcomes for core outcome sets. The methods for social network analysis, with their basis in network theory,[9] have increasingly been applied to health-related research, however. Some of these applications include evaluating collaboration among researchers;[10-12] evaluating scholarly citation patterns in Alzheimer's disease;[13] and evaluating the associations between personal relationships (e.g., spouses, friends) and happiness,[14] depression,[15] food choices,[16] physical activity,[17] alcohol consumption,[18] marijuana use,[19] and smoking.[20-22]

Understanding the affinity (or repulsion) between certain outcomes that results in their co-occurrence (or not) would help identify outcomes that are considered *central* to a network of outcomes (i.e., important to the connectedness of other outcomes in a network) in clinical trials or systematic reviews. In a social network of persons, for example, *centrality* would describe the most influential person in a network, i.e., the person most important to connecting other persons in the network. Centrality is a way of considering which outcomes and groupings of outcomes have been important to researchers in the past. There are various types of centrality we can calculate. *Betweenness centrality* is a statistic that is calculated and assigned to each node that tells us the proportion of times, out of

the maximum possible, that the outcome (node) occurs in the shortest path between two other outcomes (or nodes).

The Core Outcome Measures for Effectiveness Trials (COMET) Initiative suggests that core outcome sets for clinical trials be developed first by identifying potential outcomes and then establishing consensus among stakeholders.[5] Currently, approaches to identifying potential core outcomes vary,[23] although frequency of outcome occurrence in research studies provides key information. In the current study, using Cochrane systematic reviews, we examine whether social network analysis would provide any new information on the centrality of outcomes that have been examined in a topic area. Systematic reviews synthesize multiple clinical trials, and have a key role in directly informing clinical practice guidelines and healthcare policy. Whereas the outcomes examined in clinical trials are generally related to patient care and may be constrained by practical considerations, such as costs of outcome measurement and study power,[3] a different set of issues, such as the usefulness of outcomes to decision-makers, is likely to influence systematic reviewers. Thus, the outcomes selected by systematic reviewers may encompass both clinical and policy-relevant outcomes.

## OBJECTIVES

We used social network analysis methods to (1) understand patterns of co-occurrence of outcomes in systematic reviews of HIV/AIDS; and (2) identify

outcomes that are central to the network of outcomes examined in systematic reviews of HIV/AIDS.

## METHODS

*Choice of topic area and systematic review sample*

We selected HIV/AIDS for this case study because of the variety of interventions and outcomes inherent to the topic area. This variety lends itself nicely to the understanding of patterns of co-occurrence of outcomes. In addition, a recent systematic review [23] identified only one core outcome set in the area of HIV/AIDS, specifically for interventions addressing prevention of mother-to-child transmission among breastfeeding mothers.[24]

All Cochrane systematic reviews published by the Cochrane Review Group on HIV/AIDS in the *Cochrane Database of Systematic Reviews* as of June 30, 2013 (Issue 6) were eligible. For completed systematic reviews, we included the most recent version. For ongoing systematic reviews, we included the protocol (a peer-reviewed published document outlining the planned methods [including outcomes to be examined]).

*Data extraction*

We designed a data extraction form using Google Forms©. One investigator (IJS) extracted the following information pertaining to each systematic review: status (protocol only or completed systematic review), year of publication, and type of interventions assessed. Two investigators (IJS and CUG) independently

extracted, from the methods section of each systematic review, information about all examined outcomes, as stated in the systematic reviews. We resolved all discrepancies through discussion.

*Categorization of outcomes*

Two investigators (IJS and CUG) independently categorized each outcome into one of 14 categories: clinical/biological, behavioral, mental/social, antiretroviral prophylaxis/treatment, health services access/uptake, knowledge, testing/counseling, adverse effects, preference/satisfaction, attitudes, economic, quality-of-life, adherence, and miscellaneous. We developed and revised the categorization system during data extraction, as needed, and we completed it before we began the social network analysis.

*Social network analysis*

While the steps of social network analyses might differ based on the objectives of the analysis and the topic area, our social network analysis comprised five steps.

*Step 1: Defining the structure of the social network*

In the first step of our social network analysis, we defined the structure of the network using the concepts of *nodes*, *ties*, and *isolates* (explained in Box 2-1). We did not include all possible outcomes in our network. Instead, we defined the nodes as outcomes that occurred in more than one systematic review, thus

reducing the number of unique nodes or outcomes in the graphs we produced. We considered two outcomes as tied if they co-occurred in two or more systematic reviews. Based on this definition, we generated an adjacency matrix detailing whether or not each pair of outcomes was tied.[28, 29]

*Social network analysis – Step 2: Graphing the social network*

In step 2, we imported the adjacency matrix into UCINET 6,[30] and used NetDraw [31] to produce the graphs, applying principles outlined by Freeman.[32] Freeman recommends first using a systematic approach to position nodes on the graph. Using NetDraw, we employed the *spring embedding* layout, an approach that positions nodes by balancing the attraction of closer nodes with the repulsion of distant nodes.[32] For our analysis, this implied that outcomes that co-occurred were positioned closer to each other than outcomes that did not co-occur. To aid visualization of social network graphs, outcomes that were isolates (i.e., those that did not co-occur with any other outcome in two or more systematic reviews) were not depicted. Once nodes were positioned, we incorporated node attributes into the graph generated for the network, as recommended by Freeman.[32] Using color coding, we incorporated information about the main category we had applied to each outcome as that outcome's attribute.

*Social network analysis – Step 3: Calculating descriptive statistics of the social network*

The third step of our social network analysis was to describe the structure of the observed network using descriptive statistics. We employed specific statistics that relate to a network's cohesion (or connectedness) and centrality (extent to which a network depends on intermediary nodes).

*Statistics* related to cohesion of the network (explained in Box 2-1):

*Density* - We interpreted density to be the observed co-occurrences of outcomes as a fraction of all possible co-occurrences. A density of zero would mean that no co-occurrences were observed and a density of one would mean that all possible co-occurrences were observed. The closer the density is to one, the larger is the proportion of outcome co-occurrence.

*Component* – Components are subsets of nodes that are all tied directly or indirectly, but have no ties to nodes outside that subset [10] The existence of multiple components would suggest that there are silos of outcomes that co-occur together, but do not co-occur with outcomes outside the silo.

*Statistics* related to centrality of the network (explained in Box 2-1):

Conceptually, there are various kinds of centrality, such as betweenness and degree centrality. We chose *betweenness* centrality because it best describes how certain nodes are intermediaries, or important in connecting other nodes in a

network.[9. 33]. Connectivity of outcomes is an important idea because core outcome sets should contain outcomes that represent a range of categories of outcomes; outcomes that contribute to this connectivity might be of increased importance for core outcome sets.

*Node betweenness centrality (NBC)* - In our analysis, a high NBC for a certain outcome would suggest that the outcome is central to the network. The normalized version of this statistic allows the node betweenness centrality of nodes to be compared across different networks.[9] We calculated normalized node betweenness centrality (nNBC) for each outcome in our overall network as well as within sub-networks by type of intervention (see step 5). For each network in our analysis, we considered the seven outcomes with the highest nNBCs to be the central outcomes of that network. We chose seven because Cochrane recommends including no more than seven main outcomes in Cochrane systematic reviews.[34] Using NetDraw, we depicted differences in nNBC across outcomes graphically by plotting each node's size as proportional to its nNBC.

In order to compare measures of centrality using betweenness and degree centrality, we also computed each outcome's normalized node degree centrality.

*Network betweenness centralization* or *global betweenness centralization* (hereafter referred to as *centralization*) – Unlike NBC and nNBC which are computed separately for each node in a network, centralization is computed for the network as a whole. Centralization scores range from

zero to one, with scores close to zero suggesting that no single node pervades the network and scores close to one suggesting that a single node pervades the network. For example, if there was only a single outcome that co-occurred with most other outcomes, centralization would be close to one.

*Social network analysis – Step 4: Sensitivity analysis*

In step 4 of our social network analysis, we evaluated the robustness of the results using sensitivity analysis. We conducted a sensitivity analysis by changing the definition of a tie from co-occurrence of outcomes in two or more systematic reviews to co-occurrence in three or more systematic reviews. We picked the cutoff of three or more systematic reviews because the cutoff of one or more systematic reviews meant including too many isolated outcomes, and the cutoff of four or more systematic reviews meant including too few outcomes to meaningfully visualize patterns of co-occurrence.

*Social network analysis – Step 5: Subgroup analysis*

Step 5 of a social network analysis was to examine the results within sub-networks (or subgroups) of the entire network. We conducted a subgroup analysis using the Cochrane Review Group on HIV/AIDS's system for classifying its systematic reviews by type of intervention: (a) therapeutics, prognostics, and diagnostics ("clinical management"); (b) biomedical prevention ("biomedical

prevention"); (c) behavioral, social, and policy prevention ("behavioral prevention"); and (d) organization and financing of health services and care ("health services").

To evaluate whether social network analysis would provide more information than a simple frequency analysis of outcomes examined in our sample of systematic reviews, we compared the seven most *central* outcomes (i.e., those with the highest nNBCs) with the seven most *frequent* outcomes (i.e., those examined in the highest percentage of systematic reviews) overall and for each sub-network.

We analyzed descriptive statistics using STATA© version 12 (College Station, TX). We graphed and analyzed all networks using the UCINET 6 [30] and NetDraw [31] statistical packages.

## RESULTS

*Characteristics of systematic reviews and outcomes*

We identified 140 eligible systematic reviews (Table 2-1), all published in the year 2008 or later, of which 99 (70.7%) were completed. Almost half of the systematic reviews belonged to the intervention subgroup of clinical management (69/140; 49.3%). Box 2-2 lists examples of interventions assessed by subgroup. The 140 systematic reviews examined a median of seven outcomes each (interquartile range [IQR] 4-11, range 1-30). Across the 140 systematic reviews,

overall, there were 1140 outcomes examined, 294 of which were unique. The most frequently-examined outcomes were all-cause mortality (68/140; 48.6%), cost/cost-effectiveness for patients (50/140; 35.7%), and CD4 count (42/140; 30.0%) (Table 2-2). Also known as T-helper cells, CD4 refers to the cluster of differentiation 4 positive T-cells.

*Categorization of outcomes*

Figure 2-1 displays the classification of the 294 outcomes into the 14 categories we defined. Most outcomes fell into one of three categories: clinical/biological (160/294; 54.4%), behavioral (51/294; 17.4%), and mental/social (17/294; 5.8%). Almost half of the outcomes (138/294; 47%) appeared in only one review.

*Social network analysis – Steps 1 and 2: Defining the structure and graphing the social network*

After defining a tie as outcome co-occurrence in two or more systematic reviews (step 1), we graphed the social network (Figure 2-2). This graph served as exploratory analysis, displaying 156 of the 294 outcomes. The remaining 138 outcomes were isolates, and therefore were not displayed. Two main clusters of outcomes were evident in the graph (clinical/biological outcomes and behavioral outcomes), though the clustering was not mutually exclusive. For example, some clinical/biological outcomes (e.g., incidence of sexually-transmitted infections

[code G9], incidence of HIV infection [code G11]) clustered with behavioral outcomes. A few outcomes (e.g., cost/cost-effectiveness for patients [code L2] and adverse events [unspecified] [code A2]) appeared to bridge the two main clusters.

*Social network analysis – Step 3: Calculating descriptive statistics of the social network*

After exploratory analysis, we re-sized each outcome on the social network graph to be proportional to its nNBC (Figure 2-3). Overall, the most central outcomes were all-cause mortality (code G25), cost/cost-effectiveness for patients (code L2), and adverse events (unspecified) (code A2) (Table 2-2).
Because the centralization of the network is closer to zero than to one (0.27), it did not provide strong evidence for the domination of the social network from all 140 systematic reviews by the most central outcome (i.e., all-cause mortality, nNBC=23.9). The network appeared dense, and there was only one component; nevertheless, the overall cohesion of the network was low (density=0.09). This can be explained by the fact that the outcomes clustered and that only a few ties between outcomes bridged those clusters (i.e., the connections, and thus density, between the clusters was low). Further, most outcomes were examined in only a few systematic reviews each (median=2, IQR=1-3), limiting the overall co-occurrence of outcomes.

*Social network analysis – Step 4: Sensitivity analysis*

When we changed the definition of a tie from co-occurrence of outcomes in two or more systematic reviews to co-occurrence in three or more systematic reviews, the number of tied outcomes dropped from 156 to 96, but the density changed negligibly (from 0.09 to 0.10) and the centralization remained unchanged (0.27). Six of the top seven central outcomes were the same, while the seventh outcome 'unprotected sex' was replaced by adherence.

*Social network analysis – Step 5: Subgroup analysis*

We identified the seven most *central* and the seven most *frequent* outcomes across all 140 systematic reviews and by intervention-defined subgroups (Table 2-2 and Figure 2-4, subgroups sorted by increasing density). The lists of outcomes differed across sub-networks. When we compared the findings of the social network analysis with the frequency analysis of outcomes in the 140 systematic reviews, we noted that certain outcomes would be missed if one considered only frequency or only centrality. For example, adverse events (specified), a patient-important outcome, was one of the most central outcomes in the overall network (nNBC=14.8), but not one of the most frequent (17/140 systematic reviews; 12.1%). Similarly, the same outcome was the most central outcome in the biomedical prevention subgroup (nNBC=33.3), but not one of the most frequent (2/20 systematic reviews; 10.0%). This suggests that while this outcome was not

very frequently used in the systematic reviews, it was important to connecting other outcomes in the networks.

When we compared each network's seven most central outcomes using betweenness and degree centrality, we noted that the outcomes so identified were generally similar. For example, in the health services subgroup, the seven most central outcomes identified using betweenness and degree centrality were the same, except that the sixth (cost/cost-effectiveness for patients) and the seventh (hospitalization) most central outcomes identified using betweenness centrality swapped positions in the list of the seven most central outcomes identified using degree centrality (data not shown).

The nNBC statistics revealed that the evidence supporting the most central outcomes in the sub-networks of clinical management and biomedical prevention was stronger than the evidence supporting the most central outcomes in the other two sub-networks. This observation is a consequence of how centralization is computed, i.e., networks with large differences in nNBC between their two most central outcomes have high centralization scores. For example, in the most centralized sub-network, clinical management (centralization=0.37, Table 2-3), the most central outcome (all-cause mortality) had considerably higher nNBC than that of the next most central outcome (38.2 vs. 13.0, Table 2-2). On the other hand, in the least centralized sub-network, health services (centralization=0.12, Table 2-3), the difference in nNBC between the two most central outcomes was not as great (14.1 vs. 8.7). While all-cause mortality was the most central outcome

in both these sub-networks, it was more central to the clinical management sub-network than the health services sub-network (nNBC 38.2 vs. 14.1, respectively, Table 2-2). The comparison of nNBCs for the same outcome in different networks is valid because of the normalized nature of this statistic.

The inverse relationship between density and centralization for the sub-networks (Table 2-3) suggests that the more dense (or cohesive) sub-networks were less centralized around their most central outcome. All sub-networks included one component each, except for the biomedical prevention sub-network (Figure 2-4b) which contained two components (one of which is tiny).

## DISCUSSION

Choice of outcomes is a critical step in study design of both clinical trials and systematic reviews, and yet, current approaches to inform that choice vary. Inconsistent use of outcomes leads to systematic reviewers being unable to compare results across studies or synthesize results in a meta-analysis, threatening the credibility and potential impact of evidence synthesis. In this case study, we draw from the toolbox of social network analysis and apply these methods to identify systematically *central* outcomes (important to the connectedness of other outcomes in the network) in systematic reviews of HIV/AIDS. Examining all 140 Cochrane systematic reviews published by the Cochrane Review Group on HIV/AIDS, we identified 294 unique outcomes overall, across four pre-specified intervention subgroups defined by Cochrane. Before beginning our social network

analysis, we grouped or categorized outcomes to manage the large number of outcomes and to examine their approximate topic coverage. Whereas other researchers might reasonably differ in how they would categorize these outcomes, the general overlap between our categorization system and the spring embedding-based clustering of outcomes in the social network graphs provided *a posteriori* support for our categorization system. Moreover, our categories incorporate, either as main categories or subcategories, each of the 15 categories used during a recent survey of outcome reporting in Cochrane systematic reviews.[35] However, future investigators might use *community detection algorithms* [36] to formally evaluate the extent of overlap between pre-defined categories of outcomes and social network analysis-identified *communities* of outcomes.

We identified the seven most central outcomes for HIV/AIDS interventions overall as well as separately for each intervention sub-network, considering outcomes with the highest values of nNBC to be central to a network. For two intervention sub-networks (clinical management and biomedical prevention) in particular, we observed some evidence supporting pervasion of the networks by their single most central outcomes. While the nNBC statistic identifies the most central outcomes in a network, the difference in nNBC between the most and second most central outcomes allow7s assessment of the extent to which the most central outcome pervades the network. High centralization and large differences in nNBC imply pervasion by that outcome. Future research should evaluate whether a minimum nNBC cutoff can be used to determine whether or not a given outcome

is central to a network. Furthermore, it is unclear whether seven outcomes is the optimal number for examining centrality. In future testing of this method, researchers may want to examine whether there is additional gain from comparing, for example, a larger number of outcomes.

*What is gained by examining centrality?*

To ensure that important outcomes are not missed, we suggest that those developing core outcome sets begin by considering both frequent and central outcomes. Social network analysis capitalizes on the underlying affinity and repulsion between outcomes, thus identifying co-occurring *central* outcomes (as opposed to *frequent* outcomes) in existing research. Central outcomes, the most important outcomes in a network, are readily visualized on social network graphs and captured statistically using the nNBC statistic. When we compared the most central with the most frequent outcomes overall or in a sub-network, we observed some differences. Although there was some overlap, there was additional information conveyed by the social network analysis that took co-occurrence into account. For example, the list of most frequent outcomes, overall, excludes the outcome of adverse events (specified), a patient-important outcome that we considered central using social network analysis. So, if *frequency* were used as the sole determinant for identifying outcomes for a core outcome set, adverse events (specified) would have been missed. Similarly, in the sub-networks, all-cause

mortality, viral load, and symptom resolution are examples of outcomes that would have been missed if frequency were solely used.

When comparing the two lists of frequent and central outcomes, we observed somewhat striking differences in the more centralized sub-networks (i.e., clinical management [centralization=0.37] and biomedical prevention [centralization=0.29]) compared with the less centralized sub-networks. For example, in the biomedical prevention sub-network, three of the seven most central outcomes were not among the seven most frequent, and vice versa. On the other hand, in the health services sub-network (the least centralized sub-network), there was only one exclusive outcome in each list. One interpretation of this observation is that the social network analysis approach might be more valuable in more centralized networks compared with less centralized networks (i.e., networks in which the overlap between central and frequent outcomes is greater). Before considering this interpretation further, other disease areas should be examined to see whether this finding can be more generally applied.

There are various types of centrality defined in social network analysis. We chose to use betweenness as a measure of centrality because it identifies outcomes as central if they are important to the connectedness of the network. Degree centrality measures the number of direct connections that each node has.[8] Eigenvector centrality is the measure of the influence of a node, or outcome, in a network.[8] Closeness centrality is related to the shortest distance between pairs of nodes.[8] While the latter three types of centrality capture important

characteristics of outcomes in a network, betweenness centrality captures a concept fundamental to developing core outcome sets, the connectedness of the entire network and how the outcomes in a network relate to one another. Future investigators may wish to explore other measures of centrality. In our data, the most central outcomes identified using betweenness and degree centrality were generally similar.

*Developing core outcome sets using social network analysis*

It can reasonably be assumed that the existence and promotion of core outcome sets would enhance the comparability of outcomes across research in a given topic area.[4] Although core outcome sets represent an attempt at standardization of outcomes examined and reported, they are developed using a variety of methods, sometimes arbitrarily. Some of the methods used include the Delphi technique, semi-structured group discussions (e.g., workshops), unstructured group discussions, literature/systematic reviews, and surveys.[23] The use of formal processes has been recommended by the COMET Initiative,[23] the Patient-Reported Outcomes Measurement Information System (PROMIS),[37] and the Outcome Measures in Rheumatology (OMERACT) Initiative.[38]

While it is too early to decide on an optimal method for develop core outcome sets, we believe social network analysis can contribute a perspective that adds to traditional frequency, consensus, and survey methods. Central outcomes should not automatically be considered as core outcomes. Our study does not test

various methods for developing core outcome sets; rather, we explore a method (social network analysis) that could contribute by identifying potential outcomes (central outcomes) for core outcome sets in a single topic area, HIV/AIDS. Because we have demonstrated that some central outcomes would be missed if one were only to consider the most frequent outcomes, for the topic area of HIV/AIDS, we suggest incorporation of this perspective would add value to the process of core outcome set development. Once the list of potential outcomes is obtained from both frequent and central outcomes, assuming involvement of all stakeholders, the next step should be to develop core outcome sets. A critical component of core outcome set development is the assurance that all stakeholders, including patients, have a voice in the process.

Because our social network analysis demonstrated different central outcomes across the intervention sub-networks, we believe that tailoring core outcome sets by intervention approach is beneficial. Even in the instance where the most central outcome for two sub-networks was the same (i.e., all-cause mortality in the sub-networks of clinical management and health services), this outcome was more central to the former sub-network than the latter (nNBC 38.2 vs. 14.1). This suggests stronger evidence for including all-cause mortality in a core outcome set for clinical management than for health services in HIV/AIDS.

Another aspect related to the sub-networks in our study that bears discussion is the fact that the outcomes in the sub-networks were obtained from Cochrane-defined subgroups of HIV/AIDS reviews. In most subgroups, there were

both different types of interventions as well as different types of populations. In the overall network, the different types of interventions included both preventive and therapeutic interventions. Those developing core outcome sets in a topic area should consider whether separate core outcome sets are warranted by type of intervention, type of population, or both. The networks of outcomes would need to be constructed accordingly. The networks we constructed reflect the Cochrane Review Group on HIV/AIDS's classification of its systematic reviews.

*Developing core outcome sets starting with systematic reviews*

Systematic reviews are an excellent starting point for identifying central outcomes for core outcome sets. Existing systematic reviews in a topic area, when considered together, appraise a large portion of the evidence, much of it from clinical trials. Although it bears further investigation, deriving central outcomes from systematic reviews is potentially useful for designing clinical trials relevant for systematic reviews and decision-making. Another argument for using systematic reviews to help develop a core outcome set, is that it is recommended that systematic reviews (e.g., Cochrane systematic reviews [34]) include patient representatives in the process. If followed, this recommendation allows for broader input on outcome inclusion. Further, systematic reviews of intervention effectiveness pre-specify outcomes and indicate where these outcomes are missing from clinical trial reports.[3, 39, 40] For example, the choice of outcomes in a clinical trial may be related to what data can be gathered easily (i.e., interim

outcomes), and may not always address the questions that need to be answered. An example of this is the focus of glaucoma clinical trials on intraocular pressure (a risk factor for glaucoma) instead of visual field, the patient-important outcome that influence's visual function.[41]

Cochrane systematic reviews use fairly standardized processes within review groups. For approximately 36% of Cochrane review groups, the review group's editorial team makes decisions about outcome selection and the relative importance of outcomes for all reviews under the group's purview.[42] In the area of HIV/AIDS, however, Cochrane systematic review authors propose a list of outcomes to be examined in the systematic review, with additions and deletions suggested by the Review Group's editorial team and peer reviewers, which often include patients and other stakeholders. As far as we know, there is only one core outcome set related to the area of HIV/AIDS, and it is specifically related to prevention of mother-to-child transmission of HIV among breastfeeding mothers.[24] Thus, the central outcomes we identified likely reflect the preferences of the larger community of Cochrane systematic review authors in HIV/AIDS, rather than just those of the Review Group. However, in smaller topic areas where the same authors might contribute to a large proportion of the systematic reviews, it is possible that the central outcomes identified might be greatly influenced by those authors' preferences, rather than the larger community of stakeholders in that topic area.

It is possible, however, that systematic reviews are not the best source materials for developing a core outcome set. Systematic reviews have been shown to sometimes miss important outcomes reported in clinical trials.[40] Therefore, the number of outcomes examined in clinical trials included in the systematic reviews we examined might be greater than the number of outcomes examined in the systematic reviews alone. Although Cochrane recommends that systematic reviewers examine outcomes without considering clinical trial outcomes,[34] we believe that future studies should examine the central outcomes of networks developed using the two different sources of information (clinical trials and systematic reviews). The amount of overlap between the two networks would help address whether those developing core outcome sets should focus on outcomes examined in systematic reviews, clinical trials, or both.

*Conclusions*

We believe that the novel application of social network analysis methods to identify outcomes that are *central* to network of outcomes in HIV/AIDS using Cochrane systematic reviews provides important information needed for the development of core outcome sets. Social network analysis can uncover co-occurrence patterns of HIV/AIDS outcomes likely not discernable using simple frequency counts and other currently used methods. We identified seven of the most central outcomes across all interventions as well as for intervention sub-networks. While there was some overlap, the outcomes identified using methods

to identify co-occurrence and centrality were different from those identified using frequency of occurrence alone. Although our results are preliminary, the methods appear feasible and deserve further study.

## ACKNOWLEDGEMENTS

We thank Evan Mayo-Wilson, MPA, DPhil; Swaroop Vedula, MBBS, MPH, PhD; and Stanley Wasserman, PhD for comments on previous versions of this manuscript. We thank Xinggang Liu, PhD for assistance with the social network analysis plots; and Xue Wang, MBBS, MSPH for assistance with pilot-testing the data abstraction form. This work was funded by the National Institutes of Health grant number EY020140, the Cochrane Eyes and Vision Group, and the Johns Hopkins Center for Global Health.

## REFERENCES


1. Meinert CL. Clinical trials dictionary: Terminology and usage recommendations. 2nd edition. Wiley. Hoboken, NJ.

2. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database–update and key issues. *N Engl J Med* 2011; 364(9): 852-60.

3. Saldanha IJ, Dickersin K, Wang X, Li T. Outcomes in Cochrane systematic reviews addressing four common eye conditions: An evaluation of completeness and comparability. *PLOS One* 2014; 9(10): e109400.

4. COMET Initiative. COMET Initiative. Available online at: www.comet-initiative.org/. Last accessed May 28, 2015.

5. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012; 13(132). doi: 10.1186/1745-6215-13-132.

6. Medical Research Council Network Hubs for Trails Methodology Research. COMET Initiative. Available online at: http://www.methodologyhubs.mrc.ac.uk/upcoming_workshops/comet_initiative.aspx. Last accessed May 28, 2015.

7. National Institutes for Health Research. COMET database launched to encourage use of core outcome sets in health research. Available online at: http://www.rds.nihr.ac.uk/latest-news/comet-database-launched-to-encourage-use-of-core-outcome-sets-in-health-research/. Last accessed May 28, 2015.

8. Wasserman, S. & Faust, K. (1994). Social Network Analysis Methods and Applications. Cambridge: Cambridge University Press.

9. Freeman, L.C., 1979. Centrality in networks: I. Conceptual clarification. Social Networks 1, 215–39.

10. Eblen M, Fabsitz RR, Olson JL, Pearson K, Pool LR, Puggal M, et al. Social network analysis comparing researcher collaborations in two cardiovascular cohort studies. *Research Evaluation* 2012; 21: 392-405.

11. Newman ME. Coauthorship networks and patterns of scientific collaboration. *Proc Natl Acad Sci USA* 2004;101 Suppl 1:5200-5.

12. Hughes ME, Peeler J, Hogenesch JB, Trojanowski JQ. The growth and impact of Alzheimer disease centers as measured by social network analysis. *JAMA Neurol* 2014; 71(4):412-20.

13. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. *BMJ.* 2009 Jul 20; 339:b2680. doi: 10.1136/bmj.b2680.

14. Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ*. 2008 Dec 4;337:a2338. doi: 10.1136/bmj.a2338.

15. Rosenquist JN, Fowler JH, Christakis NA. Social network determinants of depression. *Mol Psychiatry* 2011; 16(3): 273-81.

16. Pachucki MA, Jacques PF, Christakis NA. Social network concordance in food choice among spouses, friends, and siblings. *Am J Public Health* 2011; 101(11): 2170-7.

17. de la Haye K, Robins G, Mohr P, Wilson C. How physical activity shapes, and is shaped by, adolescent friendships. *Soc Sci Med* 2011; 73(5): 719-28.

18. Rosenquist JN, Murabito J, Fowler JH, Christakis NA. The spread of alcohol consumption behavior in a large social network. *Ann Intern Med* 2010; 152(7): 426-33.

19. Tucker JS, de la Haye K, Kennedy DP, Green HD, Pollard MS. Peer influence on marijuana use in different types of friendships. *J Adolesc Health* 2014; 54(1): 67-73.

20. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *N Engl J Med* 2008; 358(21): 2249-58.

21. Valente TW, Unger JB, Johnson CA. Do popular students smoke? The association between popularity and smoking among middle school students. *J Adolesc Health* 2005; 27(4): 323-9.

22. Green HD Jr, Horta M, de la Haye K, Tucker JS, Kennedy DR, Pollard M. Peer influence and selection processes in adolescent smoking behavior: a comparative study. *Nicotine Tob Res* 2013; 15(2): 534-41.

23. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, Williamson PR. Choosing health outcomes for comparative effectiveness research: A systematic review. *PLOS One* 2014. DOI: 10.1371/journal.pone.0099111.

24. Alioum A, Dabis F, Dequae-Merchadou L, Haverkamp G, Hudgens M, Hughes J, Karon J, Leroy V, Newell ML, Richardson B, Weverling GJ. *Stat Med* 2001; 15;20(23):3539-56.

25. Valente TW. (2010). Social Networks and Health: Models, Methods, and Applications. New York: Oxford University Press.

26. Carrington PJ, Scott J, Wasserman AS (2005). Models and methods in social network analysis. Cambridge University Press, New York, NY.

27. Borgatti SP. Centrality and AIDS. *Connections* 1995; 18 (1): 112–114.

28. Breiger RL. The duality of persons and groups. *Social Forces* 1974; 53(2): 181-90.

29. Everett MG, Borgatti SP. The dual-projection approach for two-mode networks. *Social Networks* 2013; 35(2): 204-10.

30. Borgatti SP, Everett MG, Freeman LC. Ucinet for Windows: Software for social network analysis. 2002; Harvard, MA: Analytic Technologies.

31. Borgatti SP. NetDraw software for network visualization. 2002; Lexington KY: Analytic Technologies.

32. Freeman L.C. (2005). Graphic Techniques for exploring social network data. In: Carrington PJ, Scott J, Wasserman S (eds). Models and Methods in Social Network Analysis. New York: Cambridge University Press, pp. 248-269.

33. Bonacich, P. Power and centrality: a family of measures. *Am J Sociol* 1987; 92; 1170–82.

34. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available online at: www.cochrane-handbook.org. Last accessed May 28, 2015.

35. Smith V, Clarke M, Williamson P, Gargon E. Survey of new 2007 and 2011 Cochrane reviews found 37% of prespecified outcomes not reported. *J Clin Epidemiol* 2015; 68(3):237-45.

36. Lancichinetti A, Fortunato S. Community detection algorithms: a comparative analysis. Physical review 2009; E 80, 056117

37. PROMIS. About PROMIS. Available online at: http://nihpromis.org. Last accessed May 28, 2015.

38. Tugwell P, Boers M, Brooks P, Simon L, Strand, V, Idzerda L. OMERACT: An international initiative to improve outcome measurement in rheumatology. *Trials* 2007; 26; 8:38.

39. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, Williamson PR. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010; 340:c365. doi: 10.1136/bmj.c365.

40. Ioannidis JP, Horbar JD, Ovelman CM, Brosseau Y, Thorlund K, Buus-Frank ME, Mills EJ, Soll RF. Completeness of main outcomes across randomized trials in entire discipline: survey of chronic lung disease outcomes in preterm infants. *BMJ* 2015; 350:h72. doi: 10.1136/bmj.h72.

41. Saldanha IJ, Dickersin K, Wang X, Li T. Outcomes in Cochrane systematic reviews addressing four common eye conditions: An evaluation of completeness and comparability. *PLoS One* 2014; 9(10):e109400. doi: 10.1371/journal.pone.0109400. eCollection 2014.

42. Kirkham JJ, Gargon E, Clarke M, Williamson PR. Can a core outcome set improve the quality of systematic reviews? – a survey of the Co-ordinating Editors of Cochrane Review Groups. *Trials* 2013; 14:21. doi: 10.1186/1745-6215-14-21.

## BOXES

Box 2-1 – Explanation for relevant social network analysis concepts and statistics

| Concept/Statistic | Explanation |
|---|---|
| **Concepts related to *structure*** | |
| *Node* | Each individual actor in a network. |
| *Tie* or *Edge* | Underlying connection between individual nodes, represented by a line connecting two nodes on a social network graph. |
| *Isolate* | A node that is not tied to any other node in a network. |
| **Statistics related to *cohesion*** | |
| *Density* | The number of ties in a network as a fraction of the total number of ties possible.[25] |
| *Component* | A subset of nodes that are all tied directly or indirectly, but have no ties to nodes outside that subset.[10] |
| **Statistics related to *centrality*** | |
| *Geodesic distance* | The shortest path linking a given pair of nodes in a network via intermediate nodes.[8] Note that there may be multiple geodesic distances between a pair of nodes. |
| *Node betweenness centrality (NBC)* | This statistic is calculated for <u>each node in the network</u>. First, consider all geodesic distances connecting nodes 'j' and 'k' in an entire network ($g_{jk}$). Next, consider the proportion of those distances that pass through a certain other node 'i' ($g_{jk(i)}$). For node 'i', its NBC refers to that proportion, summed across all pairs of nodes in the network.[9, 33] NBC for node 'i' is computed as follows:[9, 26]<br>$NBC_i = \sum_{j<k} \frac{g_{jk(i)}}{g_{jk}}$<br>where: $g_{jk}$ = number of geodesics between nodes 'j' and 'k'<br>$g_{jk(i)}$ = number of geodesics between nodes 'j' and 'k' that pass through node 'i'<br><br>Therefore, for a given node, its NBC reflects the extent to which connections between other nodes in the network rely on that node as an intermediary.[10, 27] |

| Concept/Statistic | Explanation |
|---|---|
| *Normalized node betweenness centrality (nNBC)* | This statistic is also calculated for each node in the network. The nNBC is a normalized version of the node betweenness centrality, and is calculated by dividing the node betweenness centrality by the number of pairs of nodes in a network.[10] |
| *Network betweenness centralization* or *global betweenness centralization ("Centralization")* | This statistic applies to the network as a whole. Centralization measures the extent to which the network is centered around the most central node (i.e., the node with the highest NBC).[9, 10] Centralization is computed as follows:[9, 26]<br>$Centralization = \frac{\sum_{i=1}^{g}[NBC* - NBC_i]}{\max \sum_{i=1}^{g}[NBC* - NBC_i]}$<br>where: NBC* = largest individual NBC in the network<br>$NBC_i$ = NBC of node i<br>Therefore, the centralization simply is the sum of the difference in NBC between the most central node in the network and each of the other nodes, normalized by the maximum possible sum of the differences had the network been one where all connections depended on a single node.[26] |

Box 2-2 – Examples of interventions addressed in included Cochrane systematic reviews, by type of intervention

**Clinical management**

- Abacavir-based triple nucleoside regimens for maintenance therapy in patients with HIV
- Stavudine, lamivudine and nevirapine combination therapy for treatment of HIV infection and AIDS in adults
- Effectiveness of antiretroviral therapy in HIV-infected children under 2 years of age
- Herbal medicines for treating HIV infection and AIDS
- Topical treatments for HIV-related oral ulcers

**Biomedical prevention**

- Nonoxynol-9 for preventing vaginal acquisition of HIV infection by women from men
- Antiretroviral post-exposure prophylaxis (PEP) for occupational HIV exposure
- Interventions for preventing late postnatal mother-to-child transmission of HIV
- Sperm washing to prevent HIV transmission from HIV-infected men but allowing conception in sero-discordant couples

**Behavioral prevention**

- Behavioral interventions to promote condom use among women living with HIV
- Behavioral interventions for preventing HIV infection in homeless or unstably-housed adults
- Behavioral interventions to reduce risk for sexual transmission of HIV among men who have sex with men
- Interventions for reduction of stigma in people with HIV/AIDS
- Male circumcision for prevention of homosexual acquisition of HIV in men

**Health services**

- Home-based care for reducing morbidity and mortality in people infected with HIV/AIDS
- Integration of HIV/AIDS services with maternal, neonatal and child health, nutrition, and family planning services
- Setting and organization of care for persons living with HIV/AIDS
- Interventions to improve adherence to antiretroviral therapy in children with HIV infection
- Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection

## TABLES

Table 2-1 – Characteristics of 140 Cochrane systematic reviews examined

| Systematic review characteristic | Systematic reviews (N=140) n (%) |
|---|---|
| Year of publication | |
| 2008 | 6 (4.3) |
| 2009 | 34 (24.3) |
| 2010 | 18 (12.9) |
| 2011 | 35 (25.0) |
| 2012 | 31 (22.1) |
| 2013 | 16 (11.4) |
| Status | |
| Protocol* | 41 (29.3) |
| Completed review | 99 (70.7) |
| Type of intervention assessed | |
| Clinical management | 69 (49.3) |
| Biomedical prevention | 20 (14.3) |
| Behavioral prevention | 28 (20.0) |
| Health services | 23 (16.4) |

*A protocol is a peer-reviewed published document outlining the planned methods of a Cochrane systematic review [including outcomes to be examined]).

Table 2-2: Comparison of seven most central and seven most frequent outcomes, for all interventions and by type of intervention

| Type of intervention (number of systematic reviews) | Most central outcomes* | | Most frequent outcomes* | |
|---|---|---|---|---|
| | Outcome *(outcome code)*** | nNBC**** | Outcome *(outcome code)*** | % of systematic reviews |
| All (n=140) | All-cause mortality *(G25)* | 23.9 | All-cause mortality *(G25)* | 48.6 |
| | Cost/cost-effectiveness for patients (*L2*) | 16.4 | Cost/cost-effectiveness for patients (*L2*) | 35.7 |
| | **Adverse events (specified) *(A3)*** | 14.8 | CD4 count *(G4)* | 30.0 |
| | Quality-of-life *(K2)* | 10.6 | Quality-of-life *(K2)* | 27.9 |
| | Acquisition/incidence of HIV *(G11)* | 9.1 | Acquisition/incidence of HIV *(G11)* | 18.6 |
| | CD4 count *(G4)* | 6.5 | **Adherence *(J2)*** | 18.6 |
| | Unprotected sex (type unspecified) *(B4)* | 6.5 | Unprotected sex (type unspecified) *(B4)* | 17.9 |
| Clinical management (n=69) | All-cause mortality *(G25)* | 38.2 | All-cause mortality *(G25)* | 72.5 |
| | Adverse events (unspecified) *(A2)* | 13.0 | Adverse events (unspecified) *(A2)* | 52.1 |
| | Adverse events (specified) *(A3)* | 13.0 | CD4 count *(G4)* | 40.6 |
| | Quality-of-life *(K2)* | 9.3 | **Major/severe/serious adverse events *(A4)*** | 31.9 |
| | **Viral load *(G5)*** | 8.8 | Quality-of-life *(K2)* | 31.9 |
| | CD4 count *(G4)* | 4.5 | Adverse events (specified) *(A3)* | 21.7 |
| | **Symptom resolution *(GO15)*** | 3.5 | **AIDS-defining illness/event *(G23)*** | 20.3 |
| Biomedical prevention (n=20) | **Adverse events (specified) *(A3)*** | 33.3 | Acquisition/incidence of HIV *(G11)* | 60.0 |
| | Acquisition/incidence of HIV *(G11)* | 24.3 | **Adverse events (unspecified) *(A2)*** | 45.0 |
| | Mother-to-child transmission of HIV *(G18)* | 22.7 | Mother-to-child transmission of HIV *(G18)* | 35.0 |
| | **All-cause mortality *(G25)*** | 16.2 | Postpartum morbidity *(G16)* | 20.0 |
| | Postpartum morbidity *(G16)* | 11.2 | Neonatal/infant mortality *(G17)* | 20.0 |
| | Neonatal/infant mortality *(G17)* | 0.6 | **Stillbirth *(GO104)*** | 15.0 |
| | **Premature delivery *(GO103)*** | 0.6 | **Major/severe/serious adverse events *(A4)*** | 15.0 |

| Type of intervention (number of systematic reviews) | Most central outcomes* | | Most frequent outcomes* | |
|---|---|---|---|---|
| | Outcome *(outcome code)*** | nNBC**** | Outcome *(outcome code)*** | % of systematic reviews |
| Behavioral prevention (n=28) | **Self-efficacy/empowerment *(E2)*** | 14.9 | Acquisition/incidence of HIV *(G11)* | 46.4 |
| | Cost/cost-effectiveness for patients (*L2*) | 12.0 | Acquisition/incidence of STIs *(G9)* | 46.4 |
| | **Uptake of HIV testing *(D2)*** | 9.6 | **Number of sexual partners *(B8)*** | 28.6 |
| | Acquisition/incidence of STIs *(G9)* | 7.5 | **Prevalence of HIV *(G13)*** | 28.6 |
| | Acquisition/incidence of HIV *(G11)* | 7.5 | Cost/cost-effectiveness for patients (*L2*) | 28.6 |
| | Condom use (male condoms) *(BO35)* | 7.3 | Condom use (male condoms) *(BO35)* | 28.6 |
| | Condom use (female condoms) *(BO36)* | 7.3 | Condom use (female condoms) *(BO36)* | 28.6 |
| Health services (n=23) | All-cause mortality *(G25)* | 14.1 | All-cause mortality *(G25)* | 65.2 |
| | CD4 count *(G4)* | 8.7 | Quality-of-life *(K2)* | 56.5 |
| | Quality-of-life *(K2)* | 7.6 | CD4 count *(G4)* | 56.5 |
| | Progression to AIDS *(G22)* | 6.4 | Viral load *(G5)* | 47.8 |
| | Viral load *(G5)* | 5.8 | Progression to AIDS *(G22)* | 43.5 |
| | Cost/cost-effectiveness for patients *(L2)* | 4.6 | Cost/cost-effectiveness for patients *(L2)* | 30.4 |
| | **Hospitalization *(G20)*** | 1.7 | **Adherence *(J2)*** | 30.4 |

*Within each type of intervention, outcomes in **bold text** are exclusive (i.e., among the seven most central outcomes in a group of systematic reviews, but are NOT among the seven most frequent outcomes, or vice versa).

** Alphanumeric outcome codes are provided here to depict categories of each outcome and facilitate comparisons with Figures 2-1 through 2-4. Each alphabet denotes a separate outcome category.

*** nNBC, the normalized node betweenness centrality, reflects the extent to which connections between the other nodes (outcomes) in the network rely on that node (outcome) as an intermediary. Note that the nNBC is not a percentage.

Table 2-3: Social network statistics for all interventions and by type of intervention (sorted by increasing network density)

| Type of intervention | Number of systematic reviews | Number of outcomes | Number of components | Density | Network betweenness centralization ("centralization") |
|---|---|---|---|---|---|
| All interventions | 140 | 156 | 1 | 0.09 | 0.27 |
| Clinical management | 69 | 58 | 1 | 0.17 | 0.37 |
| Biomedical prevention | 20 | 22 | 2 | 0.22 | 0.29 |
| Behavioral prevention | 28 | 39 | 1 | 0.33 | 0.13 |
| Health services | 23 | 22 | 1 | 0.52 | 0.12 |

## FIGURES

### Legend for all Figures

The alphabet before each category name refers to the category code. Each individual outcome within a category was assigned its own alphanumeric code, beginning with the category code and followed by a number.

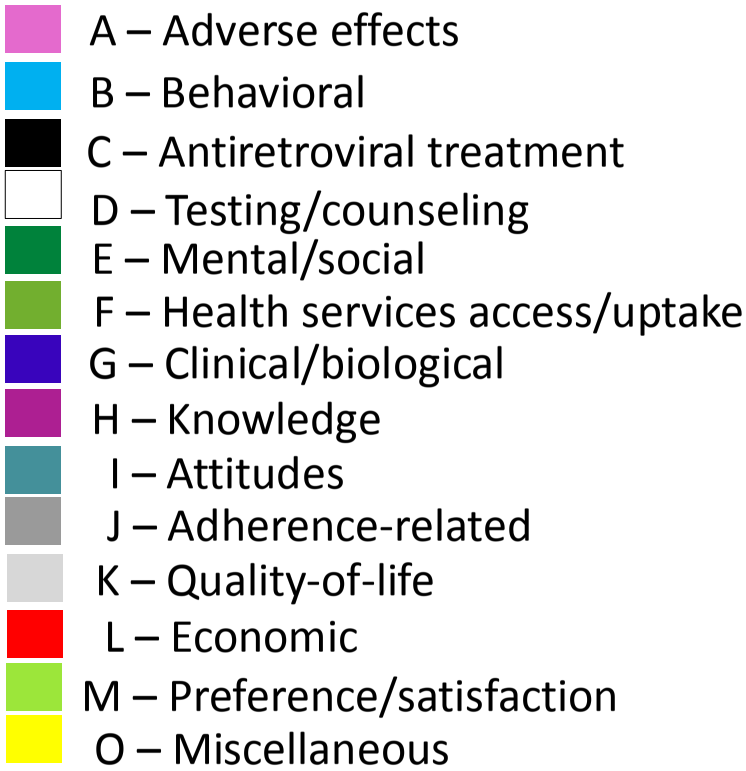
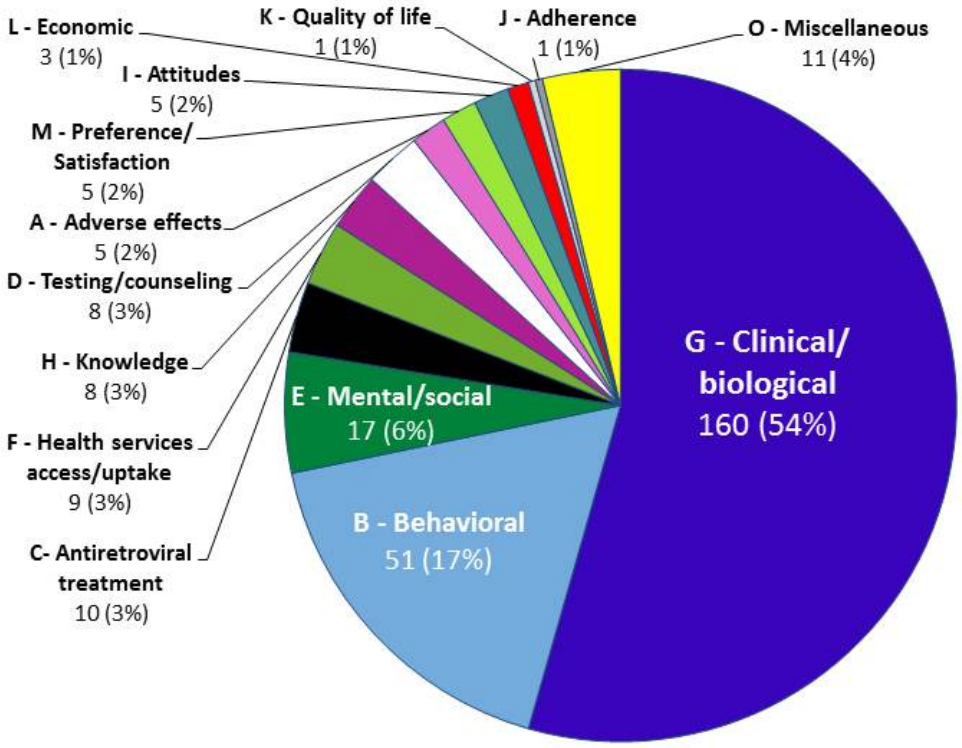


Figure 2-1: Categorization of all 294 unique outcomes into 14 categories



Examples of outcomes by category

A – Adverse events: major/severe/serious adverse events
B – Behavioral: Unprotected vaginal sex
C – Antiretroviral treatment: Switching of antiretroviral treatment
D – Testing/counseling: Pre-test counseling
E – Mental/social: Depression/depressive symptoms
F – Health services access/uptake: Utilization of healthcare
G – Clinical/biological: All-cause mortality
H – Knowledge: Condom use knowledge
I – Attitude: Sexual risk behavior attitudes
J – Adherence: Adherence
K – Quality-of-life: Quality-of-life
L – Economic: cost/cost-effectiveness for patients
M – Preference/satisfaction: Patient satisfaction with intervention
O – Miscellaneous: Illness intrusiveness

Figure 2-2: Exploratory social network analysis of 156 unique outcomes that co-occurred with at least one other outcome in two or more systematic reviews (color-coded by category of outcome). All node sizes are equal. Outcomes are labeled using their outcome codes. Alphanumeric codes are provided next to each node to depict categories of each outcome and facilitate comparisons with Table 2-2 and Figure 2-1. Each alphabet denotes a separate outcome category.

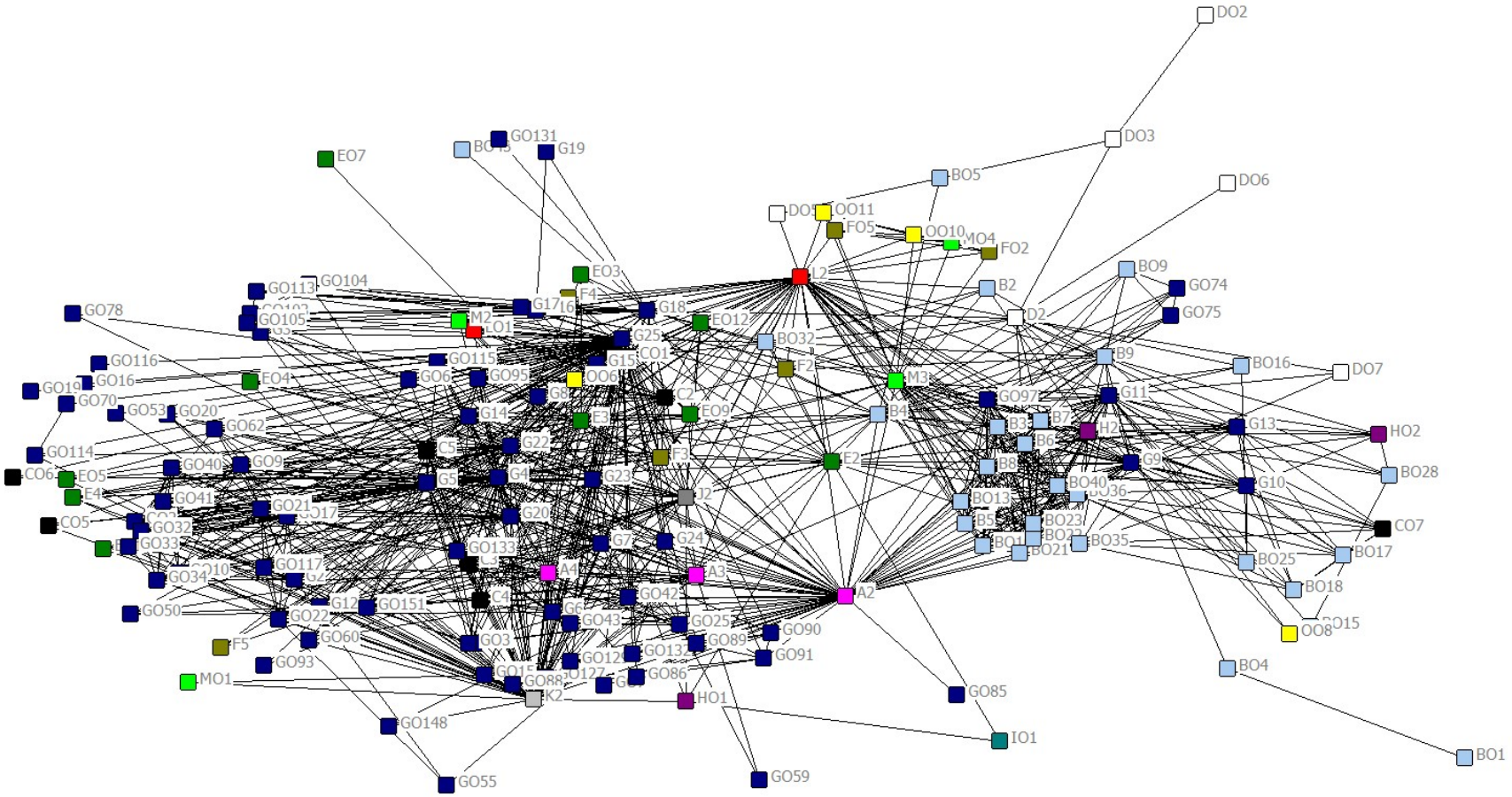


Figure 2-3: Main social network analysis of 156 unique outcomes that co-occurred with at least one other outcome in two or more systematic reviews (color-coded by category of outcome). Node size is proportional to node betweenness centrality. Outcomes are labeled using their outcome codes. Alphanumeric codes are provided next to each node to depict categories of each outcome and facilitate comparisons with Table 2-2 and Figure 2-1. Each alphabet denotes a separate outcome category.

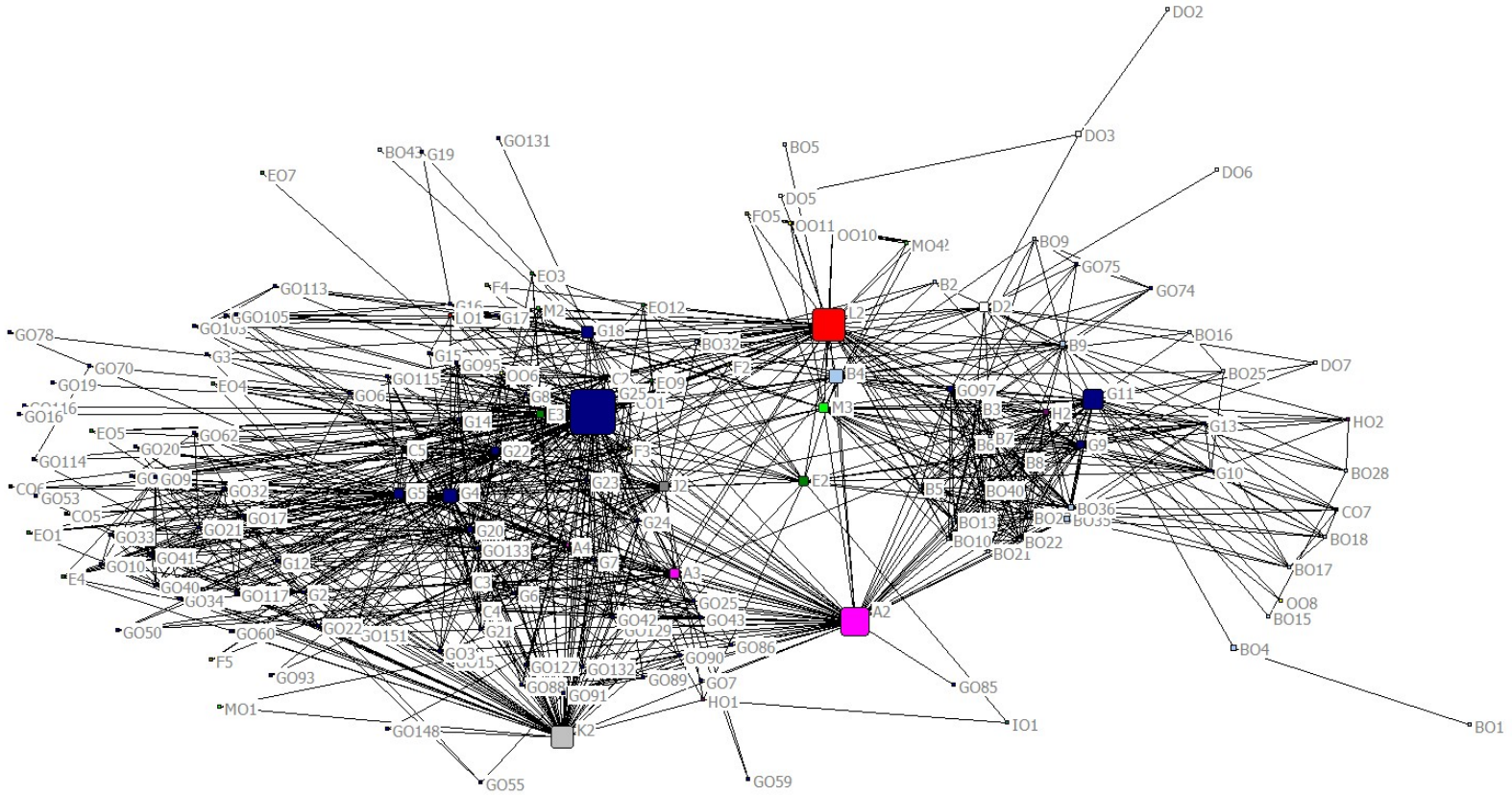


Figure 2-4: Subgroup analysis - By type of intervention (clinical management [2-4a]; biomedical prevention [2-4b]; behavioral prevention [2-4c]; and health services [2-4d]). Node size is proportional to node betweenness centrality. Outcomes are labeled using their outcome codes. Alphanumeric codes are provided next to each node to depict categories of each outcome and facilitate comparisons with Table 2-2 and Figure 2-1. Each alphabet denotes a separate outcome category.

Figure 2-4a – Sub-network of outcomes in systematic reviews assessing clinical management interventions

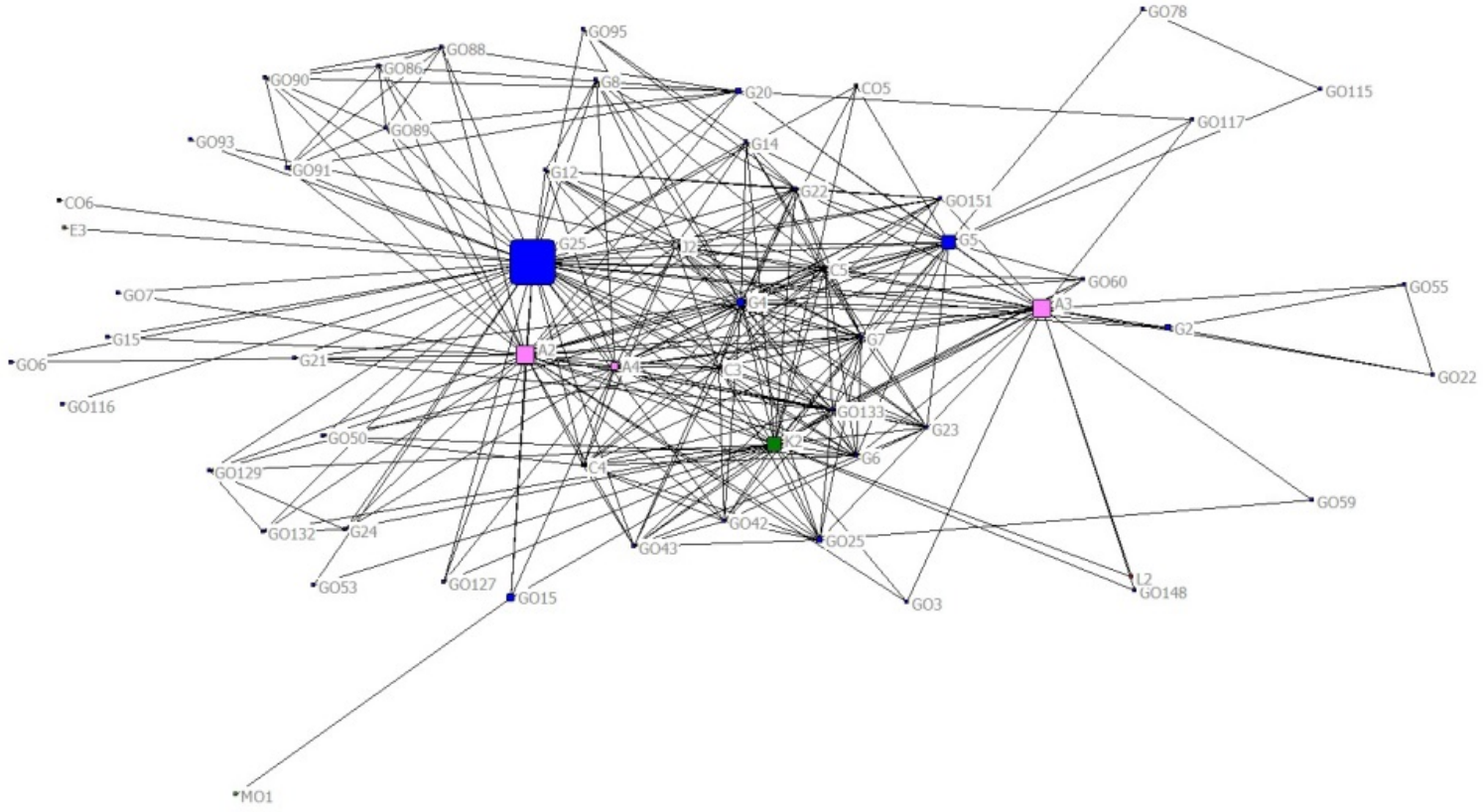


Figure 2-4b – Sub-network of outcomes in systematic reviews assessing biomedical prevention interventions

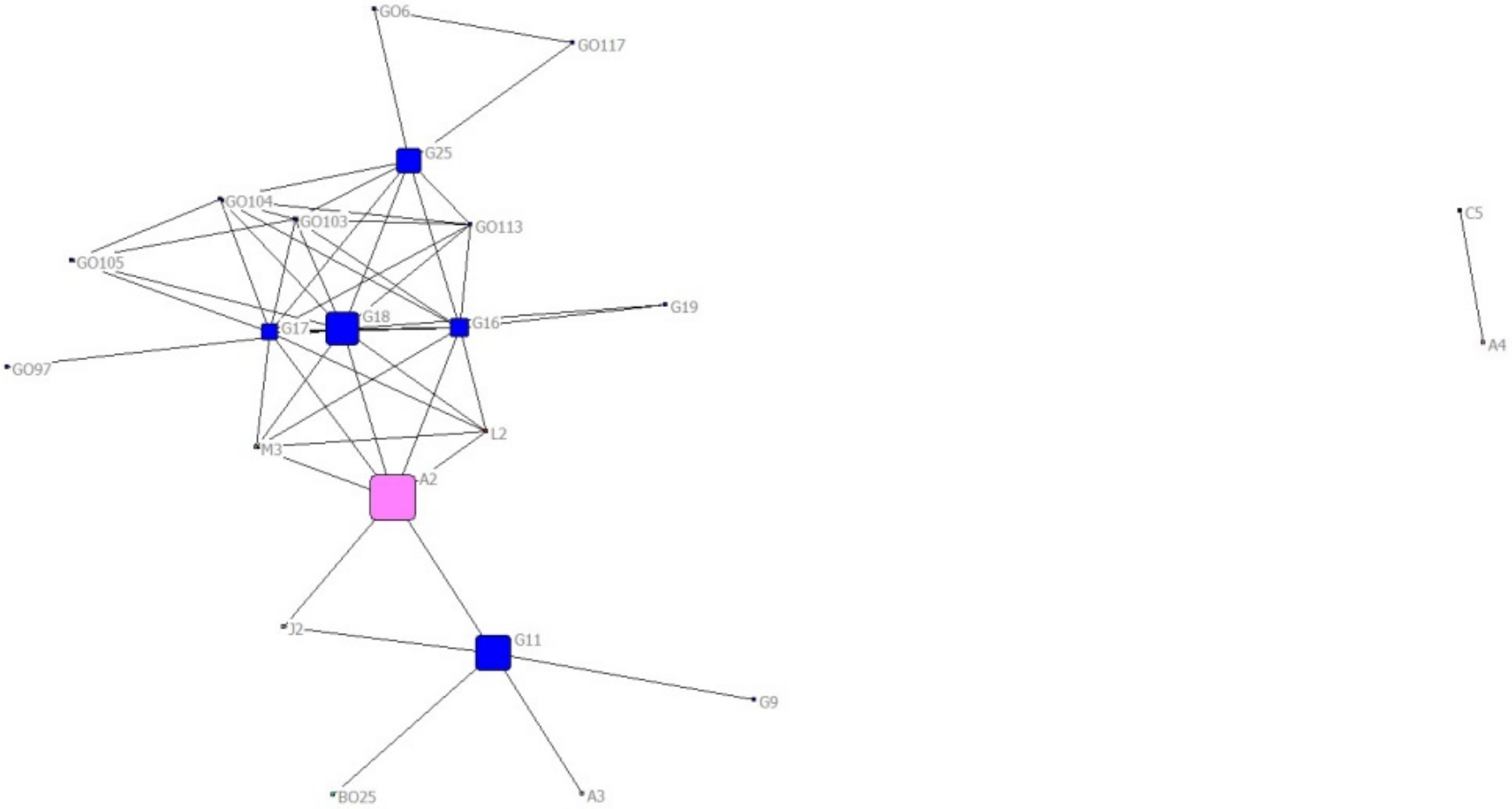


Figure 2-4c – Sub-network of outcomes in systematic reviews assessing behavioral prevention interventions

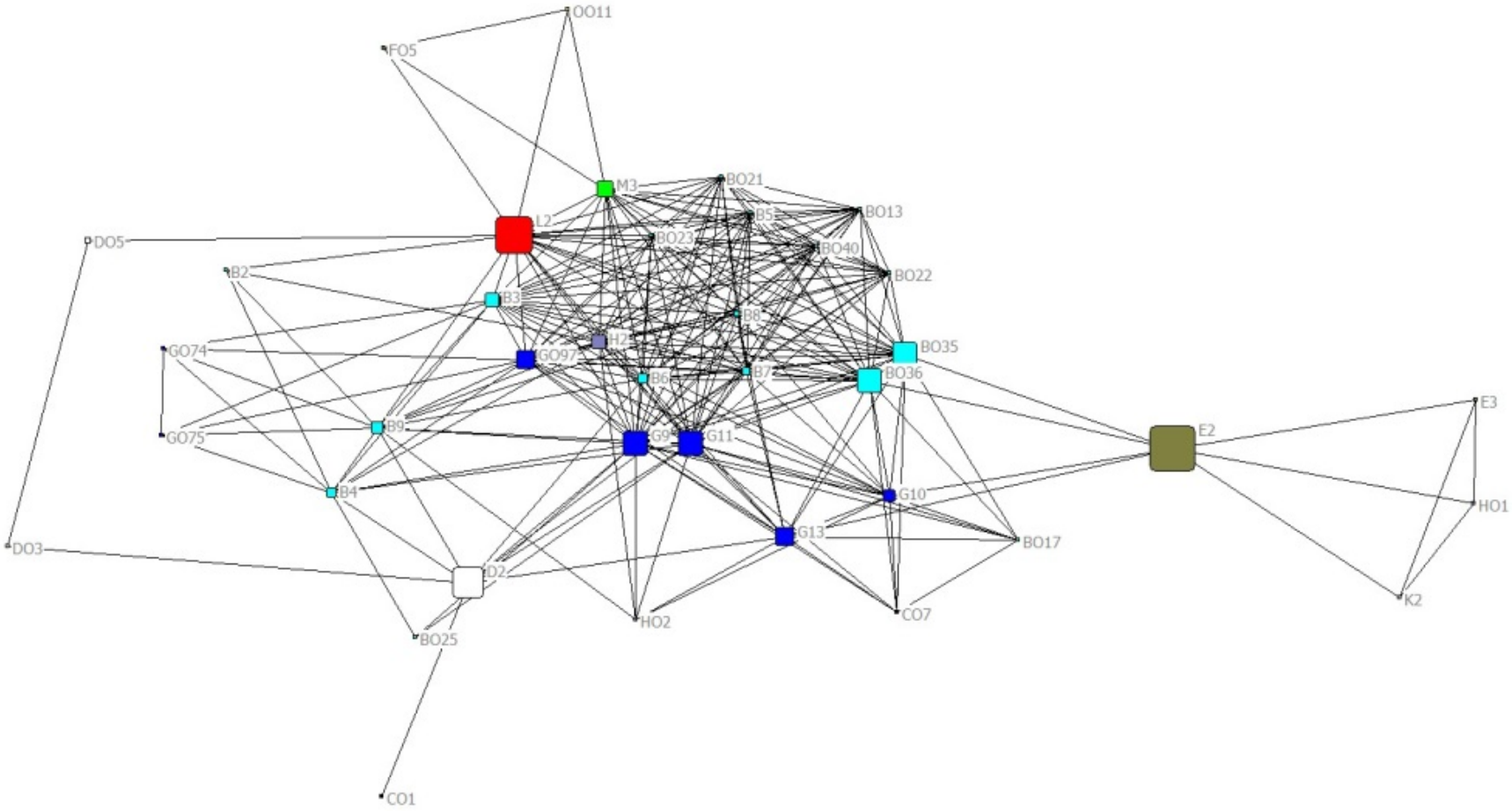
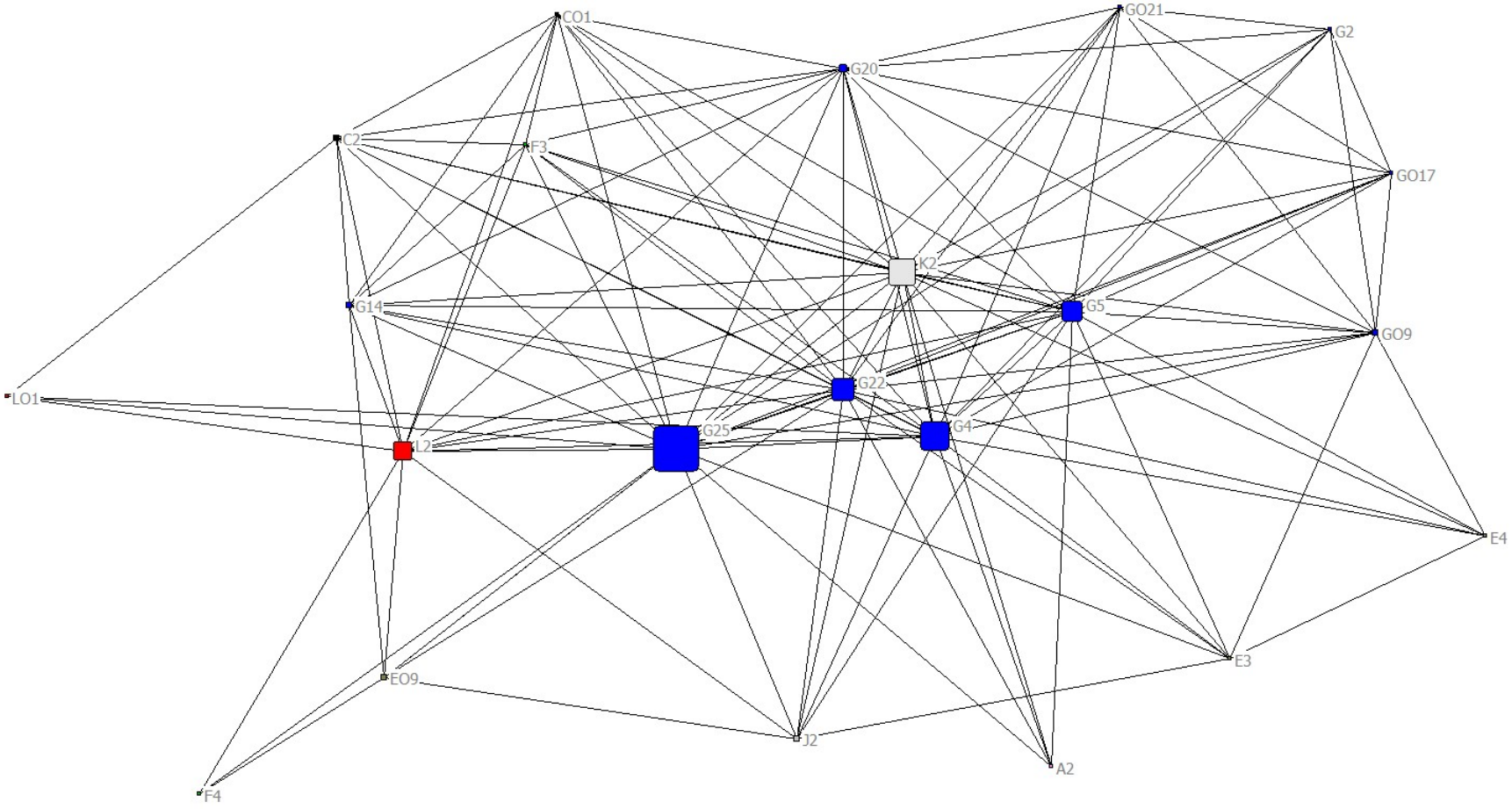


Figure 2-4d – Sub-network of outcomes in systematic reviews assessing health services interventions



# CHAPTER THREE

## MANUSCRIPT TWO

## Outcomes in Cochrane systematic reviews addressing four common eye conditions: An evaluation of completeness and comparability

**Outcomes in Cochrane systematic reviews addressing four common eye conditions: An evaluation of completeness and comparability**

Ian J Saldanha,[1] Kay Dickersin[1], Xue Wang,[1] Tianjing Li,[1]

[1] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Citation of published version:

Saldanha IJ, Dickersin K, Wang X, Li T. Outcomes in Cochrane systematic reviews addressing four common eye conditions: An evaluation of completeness and comparability. *PLoS One* 2014; 9(10): e109400.

## ABSTRACT

### Introduction

Choice of outcomes is critical for clinical trialists and systematic reviewers. It is currently unclear how systematic reviewers choose and pre-specify outcomes for systematic reviews. Our objective was to assess the completeness of pre-specification and comparability of outcomes in all Cochrane reviews addressing four common eye conditions.

### Methods

We examined protocols for all Cochrane reviews as of June 2013 that addressed glaucoma, cataract, age-related macular degeneration (AMD), and diabetic retinopathy (DR). We assessed completeness and comparability for each outcome that was named in ≥25% of protocols on those topics. We defined a completely-specified outcome as including information about five elements: *domain*, *specific measurement*, *specific metric*, *method of aggregation*, and *time-points*. For each domain, we assessed comparability in how individual elements were specified across protocols.

### Results

We identified 57 protocols addressing glaucoma (22), cataract (16), AMD (15), and DR (4). We assessed completeness and comparability for five outcome domains: quality-of-life, visual acuity, intraocular pressure, disease progression, and contrast sensitivity. Overall, these five outcome domains appeared 145 times (instances). Only 15/145 instances (10.3%) were completely specified (all five

elements) (median=three elements per outcome). Primary outcomes were more completely specified than non-primary (median=four versus two elements). Quality-of-life was least completely specified (median=one element). Due to largely incomplete outcome pre-specification, conclusive assessment of comparability in outcome usage across the various protocols per condition was not possible.

**Discussion**

Outcome pre-specification was largely incomplete; we encourage systematic reviewers to consider all five elements. This will indicate the importance of complete specification to clinical trialists, on whose work systematic reviewers depend, and will indirectly encourage comparable outcome choice to reviewers undertaking related research questions. Complete pre-specification could improve efficiency and reduce bias in data abstraction and analysis during a systematic review. Ultimately, more completely specified and comparable outcomes could make systematic reviews more useful to decision-makers.

## INTRODUCTION

In clinical trials, an outcome is an event or measure in study participants that is used to assess the effectiveness and/or safety of the intervention being studied.[1] Choosing relevant outcomes is a critical early step in the design of clinical trials and systematic reviews for a number of reasons.[2] In clinical trials, expected effect sizes on critical outcomes are used to determine sample size.[3] In addition, there is general agreement that by pre-specifying the primary and secondary outcomes and limiting the number of statistical analyses, clinical trialists reduce the likelihood of Type I error (i.e., finding a statistically significant treatment effect just by chance, in the absence of a true treatment effect) and outcome reporting bias (i.e., selectively reporting outcomes based on the strength and/or direction of the findings). Although satisfactory solutions have not yet been developed, there is growing recognition that these issues also apply to systematic reviews.[4, 5] Indeed, the Cochrane Collaboration recommends that systematic reviewers limit the number of and pre-specify all outcomes for their systematic review.[6, 7]

The process of conducting a systematic review of intervention effectiveness begins with formulating a research question, and then, finding and synthesizing the evidence from studies that address the question. In formulating the question, the systematic reviewer defines the population, intervention, comparison, and outcomes (PICO) to be examined. Studies that address the review question, typically clinical trials, should be broadly similar on the population, intervention

and comparison groups, but frequently report different outcomes from those chosen by the systematic reviewer. Clinical trialists typically measure numerous outcomes, sometimes in the hundreds.[8] It is likely that these outcomes are different from those chosen by the systematic reviewer; overlap of the chosen outcomes can vary from none to complete (Figure 3-1). In many cases, the primary outcome of interest to the systematic reviewers may not have been an outcome of interest to the clinical trialists,[9] or may not be reported clearly or consistently in the clinical trial reports or associated documents.[10] Systematic reviewers thus face an important decision: should they choose outcomes to be examined based on what they believe to be important outcomes ("systematic review author judgment") or based on what they know is reported in the relevant clinical trials ("clinical trialist judgment")?

How systematic reviewers choose outcomes and pre-specify them in systematic review protocols is currently unclear. One view is that, unlike clinical trialists, systematic reviewers should not base outcome choice on sample size/power calculations and Type I error rates. Instead, the objective of medical research should be to draw conclusions based on all sources of available evidence.[11] Systematic reviews, which are often used to inform clinical practice guidelines and policy, could and even should include all the outcomes that patients, clinicians, and policy-makers need to know about. Systematic reviews also allow elucidation of existing research gaps in a given field,[12] for example, when outcomes are not examined in trials and should be.

In our view, regardless of who chooses the outcomes to be assessed in a systematic review and how those outcomes are chosen, all outcomes need to be specified completely and clearly if they are to be of use to decision-makers.

The objective of our study was to assess the completeness of pre-specification and comparability of outcomes in all Cochrane reviews addressing four common eye conditions. Our purpose is not to hold systematic review protocols to a standard that may not have been described at the time they were published, rather it is to initiate a discussion on important questions for systematic reviewers: how should systematic reviewers choose outcomes to address in the review; how should these outcomes be reported (i.e., which elements are necessary for complete reporting) by systematic reviewers; and if outcomes are pre-specified in systematic review protocols, should these protocols be formally updated with amendments to reflect changing outcome specification?

## METHODS

*Review protocols examined*

The Cochrane Collaboration publishes and archives all its systematic review protocols, completed reviews, and updates in *The Cochrane Database of Systematic Reviews*. Protocols for systematic reviews, hereafter referred to as 'protocols', were eligible for our study if they were published by the Cochrane Eyes and Vision Group (CEVG) in *The Cochrane Database of Systematic Reviews* in or before June 2013 (Issue 6), and if they addressed any of the following four

eye conditions: glaucoma, age-related macular degeneration (AMD), cataract, and diabetic retinopathy (DR). We selected these four conditions because of their high disease burden across populations and the range of interventions addressing them.[13] For each eligible review, we identified the oldest available protocol and, when no protocol could be found for a review, we contacted CEVG editors and review authors via email to ask whether they had a copy. When these efforts were not successful, we used the most recent version of the completed review in place of the protocol.

*Five elements of a completely specified outcome*

We used an outcome definition that includes five elements: (1) the *domain* or outcome title (e.g., visual acuity); (2) the *specific measurement* or technique/instrument used to make the measurement (e.g., Snellen chart); (3) the *specific metric* or format of the outcome data from each participant that will be used for analysis (e.g., value at a time-point, change from baseline); (4) the *method of aggregation* or how data from each group will be summarized (e.g., mean, percent/proportion); and (5) the *time-points* that will be used for analysis (e.g., 3 months) (Figure 3-2). Whereas Zarin et al. specify these same elements,[8] they define the first four elements and consider time-points related to each of those four.

*Selecting outcome domains for data extraction*

Before beginning data extraction, one investigator (IS) identified all outcome domains in the Methods sections of included protocols. We then selected for data extraction those outcome domains appearing in at least 25% of eligible protocols. Then, for those eligible protocols with published completed reviews, we compared the Methods section of the protocol with the Methods section of the most recent version of the corresponding completed review, noting any differences in the specified outcome domains. We did this step to evaluate whether focusing on the protocols, some of which were published a while ago, would mean that we were assessing a different set of outcome domains than those currently being evaluated by the review authors.

*Data extraction*

We designed, tested, and finalized a data extraction form using Google Forms©. Two investigators (IS and XW) extracted data independently and resolved discrepancies through consensus or discussion with a third author (TL). We extracted data about the eye condition and year of publication of each protocol. We extracted from the Methods section the following data pertaining to each eligible outcome: type of outcome (primary, non-primary, or unclear [if not specified]) and each of the five outcome elements described earlier. For element 2, we extracted all specific measurements that were specified, or classified the specific measurement as unclear (if not specified). We classified element 3

(specific metric) into one or more of the following categories: (i) value at a time-point, (ii) time-to-event, (iii) change from baseline, and (iv) unclear (if not specified). We classified element 4 (method of aggregation) into one or more of the following categories: (i) mean, (ii) median, (iii) percent/proportion, (iv) absolute number, and (v) unclear (if not specified). For element 5, we extracted all time-points that were specified, or classified the time-points as unclear (if not specified).

*Data analysis*

We assessed the extent of completeness using the number of elements specified out of five possible, and considered an outcome specified in the Methods section as "complete" if all five elements were specified. For each outcome, we calculated median, interquartile range (IQR), and proportion of outcome elements specified. We performed Kruskal-Wallis tests for nonparametric comparisons of medians and distributions of extent of completeness by condition addressed, year of protocol publication, type of outcome, and outcome domain.

We assessed the frequency and comparability of outcome elements (i.e., similarity of categories for each element) for elements 3 and 4 across protocols addressing each of the four eye conditions. Protocols could specify more than one category for a given element. Comparability was therefore assessed as the distribution of those categories across protocols. As an example, if one protocol specified visual acuity at a time-point as well as change in visual acuity from

baseline, we counted both categories for specific metric (element 3). In another example, protocols addressing cataract and assessing the outcome of visual acuity were considered to be comparable in method of aggregation (element 4) if they all specified mean or all specified median or both. However, they would not be comparable in element 4 if some specified mean and others specified median.

Statistical significance was defined at the 5% level. All data were analyzed using STATA© version 12 (College Station, TX).

## RESULTS

*Characteristics of protocols examined*

Our search identified 57 eligible systematic reviews (Table 3-1). We were able to find protocols for 54 reviews (94.7%), and used the Methods section of completed reviews for the remaining three (5.3%). An updated protocol was published for one of the 54 protocols. Glaucoma was the most frequently addressed condition (22 protocols), followed by cataract (16 protocols), AMD (15 protocols), and DR (4 protocols). Approximately half of the protocols (29/57; 50.9%) were published between 2006 and 2010. Thirty-four protocols were associated with a completed review, the most recent version of which was published a median of five (IQR 4-8, range 0-15) years after publication of the protocol.

*Outcome domains used in protocols*

We examined five outcome domains named in at least 25% of the eligible protocols (Table 3-2): quality-of-life (47/57 protocols; 82.5%), visual acuity (47/57; 82.5%), intraocular pressure (21/57; 36.8%), disease progression (15/57; 26.3%), and contrast sensitivity (15/57; 26.3%). One protocol did not name any of these five outcome domains. For most completed systematic reviews (30/34; 88.2%), these five outcome domains were similar to what was named in their corresponding protocols. Compared to their protocols, two completed systematic reviews dropped quality-of-life while one completed review added it. One completed systematic review dropped contrast sensitivity.

*Completeness of outcome pre-specification*

Across the 57 protocols, the five most frequent outcome domains appeared 145 times ('instances'); however, only 15/145 instances (10.3%) involved complete pre-specification (i.e., where all five elements of the outcome were specified). Overall, a median of three (IQR 2-4) elements were specified per outcome (Table 3-3). Extent of completeness was not statistically significantly different by condition. Completeness of outcome specification may be better in protocols published later compared to earlier, (median of three [IQR 2-4] elements specified in 2006-2010 versus one [IQR 1-3] in 2000 or earlier), although the difference was not statistically significant (p=0.1635).

Fifty-four of 57 protocols (94.7%) specified at least one primary outcome. Among the five outcome domains evaluated in our study, at least one was a primary outcome in 48/57 (84.2%) protocols. Extent of completeness appeared to differ by outcome type, with primary outcomes being most completely specified and outcomes with type unclear being least completely specified (median four versus one respectively, p=0.0001). Intraocular pressure was the most completely specified outcome in our sample, with a median of four (IQR 3-4) elements specified (Table 3-2). Quality-of-life was least completely specified, with a median of one (IQR 1-2) element specified. The patterns of completeness of individual elements were similar across outcomes (Figure 3-3). Method of aggregation was specified least often, while domain and time-points were specified more often than other elements. The completeness of individual elements for the quality-of-life outcome was less than for other outcomes, overall. Although intraocular pressure was the most completely specified outcome, only 24% of protocols assessing it specified the specific measurement. Patterns of completeness of individual outcome elements also appeared to be similar across conditions, except for outcomes in DR protocols, where there were only four protocols and so the percentages are unlikely to be reliable (Figure 3-4).

Table 3-4 provides some examples of incomplete specification of outcomes in our sample of systematic reviews.

*Comparability of outcome elements*

Table 3-5 shows the distribution of specific metrics (element 3) and methods of aggregation (element 4) across instances of usage of outcome domain, by condition. The specific metric was unclear for large proportions of individual instances (often as high as 100% for the 16 instances of usage of quality-of-life in protocols addressing glaucoma and for the four instances of usage of contrast sensitivity in protocols addressing cataract). For instances where the specific metric was specified, the most frequent specific metrics were 'value at a time-point' and 'change from baseline'.

The method of aggregation was unclear for large proportions of individual instances (often as high as 100% for the 16 instances of usage of quality-of-life in protocols addressing glaucoma and for the four instances of usage of visual acuity in protocols addressing DR). For instances where the method of aggregation was specified, the most frequent methods of aggregation were 'mean' and 'percent/proportion'.

## DISCUSSION

*Summary of main findings*

We have shown that, if outcome pre-specification in systematic review protocols is judged using recommended standards for clinical trials, then it is largely incomplete. Although completeness appears to have improved somewhat over time, on average, only three of five standard elements of an outcome were

pre-specified. Due to largely incomplete outcome pre-specification, a conclusive assessment of comparability in outcome elements across the various protocols per condition was not possible. However, we observed variation in specific metrics and methods of aggregation.

*Completeness of outcome pre-specification*

There are some reasons that might explain why outcomes were not completely specified in our study of systematic review protocols. First, although we believe complete specification of all five elements is necessary for a number of reasons, the idea is new to the systematic review community. This is demonstrated by the fact that the Cochrane Handbook states only that the name of the outcome (equivalent to *domain* [element 1]), type of scale (equivalent to *specific measurement* [element 2]), and timing of measurement (equivalent to *time-points* [element 5]) must be pre-specified[6]; and there is no explicit mention of pre-specification of *specific metric* (element 3) or *method of aggregation* (element 4). Indeed, elements 1 and 5 were the most often-specified elements in our sample of protocols, though element 2 was frequently not specified (70% of the time) (Figure 3-3).

Another possible explanation for incomplete pre-specification of outcomes is that choice of outcomes could be influenced by the findings of (and outcomes examined in) the clinical trials that would be included in the review. We did not assess the outcomes examined at the level of the clinical trials to determine the

likelihood that this occurred, but suggest that doing so may contribute to a better understanding of how review outcomes are chosen. Are they chosen because systematic reviewers consider them the most important outcomes to examine, because they are the outcomes that have been examined in clinical trials, or both? If the review outcomes were chosen purely because they were the outcomes that have been reported in clinical trials, this is troubling because of the possibility of "meta-bias". We know, for example, that outcomes reported in clinical trials could have been selectively reported because of desirable or undesirable findings.[14, 15] By pre-specifying in the protocol the outcomes to be examined in the review, systematic reviewers minimize the potential for bias,[5, 16] and reassure readers that the choice of outcomes was not influenced by the results of individual clinical trials. That said, systematic reviewers are usually familiar with their field and *a priori* aware of potentially eligible clinical trials and/or how the outcome in question is frequently measured. Complete pre-specification also could improve efficiency in data abstraction and analysis during a systematic review.

Systematic reviewers may also anticipate potential variation in outcomes across included clinical trials, and may allow for this by pre-specifying the elements of the outcome domain of interest in broad rather than specific terms (e.g., "visual acuity" versus "change in visual acuity from baseline to 1 year, as measured using a Snellen chart"). If such variation is suspected, systematic reviewers could explicitly state that all variations of a given element(s) will be

included. This could minimize the occurrence of what Page et al. refer to as "selective inclusion" in systematic reviews.[5]

We assume that primary outcomes for both clinical trials and systematic reviews are chosen based on perceived clinical importance and/or importance to patients; and that they are usually measured and reported more thoroughly than non-primary outcomes.[17] Not surprisingly, in our study, primary outcomes were more completely specified than other outcome types. Our estimate of 94.7% protocols pre-specifying a primary outcome is somewhat higher than the 88% that has been reported as pre-specified in clinical trial protocols,[18] and this could be related to the fact that we were examining protocols entered into software that requests the domain names of the pre-specified outcomes.

In our study, the most incompletely pre-specified outcome was quality-of-life, a key patient-important outcome. This finding is concordant with other studies that have found that outcome reporting in clinical trials is a bigger problem for patient-important outcomes than other types of outcomes.[19]-[20] Further, when patient-important outcomes are not primary outcomes in clinical trials, the likelihood that reporting is complete is further reduced.[20] Our study aimed to evaluate the completeness and comparability of all outcomes, both patient-important and not.

Our recommendation is that systematic reviewers should engage in discussion about and strongly consider pre-specifying all five elements of each outcome they wish to examine. When explicit pre-specification of all five

elements of a given outcome is not possible, for example when all possible options for a given outcome element are not known or are too numerous, the systematic reviewers should enumerate all known acceptable options for each element and explicitly state that all options for that element would be accepted, or provide rationale for why it is impossible to completely pre-specify an element.

The Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocols (PRISMA-P) is currently under development.[21] We hope that the availability of reporting guidelines (including details about outcome specification) will improve the completeness of specification of outcomes. Assuming that the Cochrane Collaboration recognizes the importance of completeness of pre-specification, there are some possible ways to ensure that review authors are aware of the five elements of a completely specified outcome. First, editorial teams at Cochrane Review Groups (CRGs) should make all review authors aware of the five outcome elements early in the process (no later than the protocol development stage). Second, peer reviewers should be directed to consider whether the outcomes are completely pre-specified and not likely to have been chosen based on the strength and direction of the findings for those outcomes. Third, the Cochrane Handbook and other systematic review guidance materials, in addition to training workshops and other educational avenues, should incorporate explicit descriptions of all five outcome elements. Other organizations producing guidance on systematic review methodology (e.g., Agency for Healthcare Research and

Quality [AHRQ], the Centre for Reviews and Dissemination [CRD]) should also incorporate descriptions of the five outcome elements in their guidance materials.

Organizations such as the Cochrane Collaboration suggest limiting the number of outcomes examined in a systematic review.[6] However, in order to evaluate whether the effect of an intervention persists over time, an otherwise identical outcome (i.e., identical in the other four elements) is often measured at a number of time-points. For the purpose of counting the number of outcomes measured, we recommend that these repeated measurements be counted as one outcome regardless of the number of time-points at which the outcome is assessed.

*Comparability of outcome elements*

In the era of evidence-based medicine, decision-makers in healthcare (e.g., patients, clinicians, and policy-makers) increasingly rely on systematic reviews. It is important that decision-makers have access to high quality and up-to-date individual systematic reviews as well as are able to compare results across systematic reviews. Cochrane "overviews" (Cochrane reviews which compile evidence from related reviews of interventions into a single accessible and usable document),[6] and network meta-analyses (analyses of three or more interventions for a given condition in one meta-analysis [22, 23]) are examples of formal comparisons across systematic reviews. To better feed into these formal comparisons and clinical practice guidelines, the elements of outcomes used in the various systematic reviews addressing a given condition should be comparable. In

our study, the largely incomplete pre-specification of outcomes in protocols restricted our ability to assess comparability in outcome elements across protocols. In cases where the various elements were specified, however, we observed variation in specific metrics and methods of aggregation. An example of such variation is: one protocol pre-specified that the outcome domain of visual acuity would be measured as mean change in visual acuity (number of letters) from baseline to one year, while another protocol pre-specified that visual acuity would be measured as percent of participants with improvement in visual acuity of at least three letters at one year. While both protocols specified the same outcome domain at the same time-point, differences in the specific metric (mean change versus value at a time-point) and method of aggregation (mean versus percent) would preclude a direct comparison of the visual acuity results.

Efforts to promote comparability of outcomes across related clinical trials have led to the creation of core outcome measures within research fields.[24-26] One such effort is the Core Outcome Measures in Effectiveness Trials (COMET) Initiative,[27] whose investigators have produced guidance on methods for identifying core outcome sets.[28] Because the issue of comparability of outcomes across systematic reviews is complex, we recommend that researchers within a field (e.g., systematic reviewers, Cochrane review group editors, clinical trialists) and patients consider developing comparable outcomes across systematic reviews, adding to a core list over time as appropriate.

There are pros and cons of establishing comparability in outcomes across reviews, however. Increased comparability will likely facilitate formal comparisons across systematic reviews and development of clinical practice guidelines. In addition, decision-makers would be better able to compare more directly the effectiveness of treatment options. For example, hundreds of measurement scales (*specific measurements*) have been used to assess mental status in schizophrenia [29] and quality-of-life,[30] making comparability across clinical trials very challenging. Finally, use of comparable outcomes could discourage authors from 'cherry-picking' outcomes to be used in their studies.[31]

On the other hand, comparability across reviews is not always possible or desirable. Limiting outcomes to those used by previous researchers risks excluding an outcome that is in fact important, or authors may be compelled to include an outcome that they do not consider important. Additionally, it might not be possible to identify *a priori* all relevant outcomes and outcome elements for a rapidly evolving field or for a field with a large number of relevant outcomes.

*Availability of protocols and amendments to protocols*

We were unable to obtain 3/57 (5.3%) protocols associated with our sample of Cochrane reviews. This poses a concern for investigators conducting methodological research in systematic reviews, and for users of systematic reviews generally. Although we do not believe that relying on the Methods sections of three completed Cochrane reviews in the cases where we could not

find the protocols is likely to have influenced our findings, we believe that all protocols and previous versions of completed systematic reviews should be made available to researchers. Furthermore, an updated protocol was published for only one of the protocols we examined. The Cochrane Collaboration should consider keeping all protocols up-to-date by publishing updated versions of protocols or publishing protocol amendments for all its reviews. In this way, Cochrane review protocols would be formally amended in the same way that clinical trial protocols are amended and made available, providing an accessible audit trail. This practice will facilitate Cochrane's contribution of its protocols and updates to PROSPERO,[32, 33] an international database of prospectively registered systematic reviews.

Our focus on Cochrane reviews is both a strength and a limitation. Assuming that Cochrane reviews are among the most rigorously conducted and reported systematic reviews,[34, 35] it is likely that completeness and comparability of outcomes are higher in our sample of reviews than in other reviews. It would be useful to know how others producing systematic reviews (e.g., AHRQ, CRD, independent authors) choose and describe outcomes in their systematic reviews.

As discussed, we did not examine the individual clinical trials examined by each Cochrane review in our sample to learn more about the source of non-comparability in outcome elements. Nor did we test for empirical evidence of outcome reporting bias on the part of the systematic reviewers. Because our

assessments of completeness and comparability were based on what was reported in the protocols (and some completed reviews), it is possible that our findings were a consequence of unsatisfactory reporting and that the rationale for the outcomes chosen could not be determined without asking the systematic review authors directly.

Our study should be replicated in other disease areas and on a larger scale to assess the applicability of our findings to other fields. Although we have compared the outcomes pre-specified in the protocol with what is in the corresponding completed review's Methods section, a next step would be to compare the outcomes in the Methods with those in the Results section. This would allow a confirmation of the potential bias by systematic reviewers that has been demonstrated by Kirkham et al. using a cohort of Cochrane reviews [36] and by various investigators studying this issue in clinical trials.[14, 17, 37, 38]

*Conclusions*

We recommend that systematic review authors strongly consider pre-specifying all outcomes of interest using the five elements of a completely specified outcome (domain, specific measurement, specific metric, method of aggregation, and time-points), amending the protocol formally, as needed. We further suggest that researchers and other stakeholders, such as patients, carefully consider the pros and cons of establishing comparability in outcomes across systematic reviews addressing a given condition.

## ACKNOWLEDGEMENTS

We acknowledge the contributions of Deborah Zarin, MD; Karen Robinson, MA, PhD; Swaroop Vedula, MD, MPH, PhD; and Evan Mayo-Wilson, MSc, MPA, DPhil for their comments on previous versions of this manuscript. We also acknowledge the contributions of Michael Marrone, MPH, Kristina Lindsley, MS, and James Heyward, BA for their help with data abstraction and locating systematic review protocols for this study.

## REFERENCES


1. Meinert CL. *Clinical trials dictionary: Terminology and usage recommendations*. 2nd edition. 2012. Wiley. Hoboken, NJ

2. Institute of Medicine. *Finding what works in health care: standards for systematic reviews*. March 23, 2011. Available: http://www.iom.edu/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews.aspx. Accessed July 17, 2015.

3. Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *BMJ* 1995; 311(7013):1145-1148.

4. Bender R, Bunce C, Clarke M, Gates S, Lange S, et al. Attention should be given to multiplicity issues in systematic reviews. *J Clin Epidemiol* 2008; 61(9):857-865.

5. Page MJ, McKenzie JE, Forbes A. Many scenarios exist for selective inclusion and reporting of results in randomized trials and systematic reviews. *J Clin Epidemiol* 2013; 66(5):524-537.

6. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available: www.cochrane-handbook.org. Accessed July 17, 2015.

7. Chandler J, Churchill R, Higgins J, Tovey D. *Methodological standards for the conduct of new Cochrane Intervention Reviews*. Version 2.2. 17 December 2012. Available: http://www.editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/MECIR_conduct_standards%202.2%2017122012.pdf. Accessed July 17, 2015.

8. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database–update and key issues. *N Engl J Med* 2011; 364(9):852-860.

9. Singh S, Loke YK, Enright PL, Furberg CD. Mortality associated with tiotropium mist inhaler in patients with chronic obstructive pulmonary disease: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2011; 14; 342:d3215. doi: 10.1136/bmj.d3215.

10. Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database Syst Rev* 2012; (1):CD008965.

11. Goodman SN. (1989) Meta-analysis and evidence. *Control Clin Trials* 1989; 10(2):188-204.

12. Robinson KA, Saldanha IJ, Mckoy NA. Development of a framework to identify research gaps from systematic reviews. *J Clin Epidemiol* 2011; 64(12):1325-1330.

13. National Eye Institute. Statistics and Data. 2010. Available: http://www.nei.nih.gov/eyedata. Accessed July 17, 2015.

14. Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *N Engl J Med* 2009; 361(20):1963-1971.

15. Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, et al. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev* 2011; Issue 1. Art. No.: MR000031. DOI: 10.1002/14651858.MR000031.pub2.

16. Stewart L, Moher D, Shekelle P. Why prospective registration of systematic reviews makes sense. *Syst Rev* 2012; 1:7. doi: 10.1186/2046-4053-1-7.

17. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010; 340:c365. doi: 10.1136/bmj.c365.

18. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009; 302(9): 977-984.

19. Wieseler B, Kerekes MF, Vervolgyi V, Kohlepp P, McGauran N, et al. Impact of document type on reporting quality of clinical drug trials: a comparison registry reports, clinical study reports, and journal publications. *BMJ* 2012; 344 d8141 doi: 10.1136/bmj.d8141

20. Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervolgyi V, et al. Completeness of reporting of patient-relevant clinical trial outcomes: comparison of unpublished clinical study reports with publicly available data. *PLOS Med* 2013; 10 (10) e1001526. doi: 10.1371/journal.pmed.1001526.

21. EQUATOR Network. Reporting Guidelines under development. Available: http://www.equator-network.org/library/reporting-guidelines-under-development/#99. Accessed July 17, 2015.

22. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521):897-900.

23. Li T, Puhan M, Vedula SS, Singh S, Dickersin K for the Ad Hoc Network Meta-analysis Methods Meeting Working Group. Network meta-analysis – highly attractive and more methodological research is needed. *BMC Medicine* 2011; 9(1):79. doi: 10.1186/1741-7015-9-79.

24. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981; 47(1):207-214.

25. Tugwell P, Boers M, Brooks P, Simon LS, Strand V. OMERACT: An international initiative to improve outcome measures in rheumatology. *Trials* 2007; 8:38.

26. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, et al. Core outcome measures for chronic pain trials: IMMPACT recommendations. *Pain* 2005; 113(1-2):9-19.

27. COMET Initiative. Overview. Available: http://www.comet-initiative.org/about/overview. Accessed July 17, 2015.

28. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012; 13(132). doi: 10.1186/1745-6215-13-132.

29. Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. BMJ 1998; 317(7167):1181–1184.

30. Salek S. *Compendium of Quality of Life Instruments*. 1999. Wiley. ISBN: 0-471-98145-1

31. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007; 8:39.

32. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Sys Rev* 2012; 1:2. doi: 10.1186/2046-4053-1-2.

33. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, et al. PROSPERO at one year: an evaluation of its utility. *Sys Rev* 2013; 2:4. doi: 10.1186/2046-4053-2-4.

34. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, et al. Methodology and reports of systematic reviews and meta-analyses: A comparison of Cochrane reviews with articles published in paper-based journals. *JAMA* 1998; 280(3):278-280.

35. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med* 2007; 4(3): e78.

36. Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in prespecified outcomes during the systematic review process. *PLoS One* 2010; 5(3):e9810. doi: 10.1371/journal.pone.0009810.

37. Chan AW, Altman D. Identifying outcome reporting bias in randomized trials on PubMed: review of publications and survey of authors. *BMJ* 2005; 330(7494):753.

38. Chan AW, Hrobjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004; 291(20):2457-2465.

## TABLES

Table 3-1: Number of protocols and outcome domains by condition, year published, and whether specified as primary outcome

| Characteristic | Number (%) of protocols | Number (%) of outcomes |
|---|---|---|
| All | 57[1] (100) | 145[2] (100) |
| Condition addressed | | |
| Glaucoma | 22 (38.6) | 51 (35.2) |
| Cataract | 16 (28.1) | 35 (24.1) |
| Age-related macular degeneration (AMD) | 15 (26.3) | 47 (32.4) |
| Diabetic retinopathy (DR) | 4 (7.0) | 12 (8.3) |
| Year of protocol publication | | |
| 2000 or earlier | 6 (10.5) | 13 (9.0) |
| 2001 to 2005 | 15 (26.3) | 37 (25.5) |
| 2006 to 2010 | 29 (50.9) | 76 (52.4) |
| 2011 or later | 7 (12.3) | 19 (13.1) |
| Type of outcomes domain specified | Not applicable | |
| Outcomes specified as primary | | 48 (33.1) |
| Outcomes specified as non-primary | | 88 (60.7) |
| Type of outcome unclear | | 9 (6.2) |

[1] 54 protocols and 3 completed reviews; One protocol did not include any of the outcome domains selected for detailed data extraction

[2] 139/145 of the outcomes were described in the 54 protocols.

Table 3-2: Completeness (number of completely-specified elements out of five possible) by outcome domain

| Characteristic | Number (%) of protocols | Median (IQR) number of completely-specified elements per outcome | p-value |
|---|---|---|---|
| All | 57[1] (100) | 3.0 (2.0 - 4.0) | - |
| Outcome domain | | | |
| Quality-of-life | 47 (82.5) | 1.0 (1.0 - 2.0) | |
| Visual acuity | 47 (82.5) | 3.0 (2.0 - 4.0) | |
| Intraocular pressure | 21 (36.8) | 4.0 (3.0 - 4.0) | 0.0001 |
| Disease progression | 15 (26.3) | 3.0 (2.0 - 4.0) | |
| Contrast sensitivity | 15 (26.3) | 2.0 (1.0 - 3.0) | |

[1] One protocol did not include any of the outcome domains selected for detailed data extraction

Table 3-3: Completeness (number of completely-specified elements out of five possible) by type of protocol/outcome

| Characteristic | Median (IQR) number of completely specified elements per outcome | p-value |
|---|---|---|
| All[1] | 3.0 (2.0 - 4.0) | NA |
| Condition addressed | | |
| Glaucoma | 3.0 (2.0 - 4.0) | 0.1218 |
| Cataract | 3.0 (2.0 - 4.0) | |
| Age-related macular degeneration (AMD) | 2.0 (1.0 - 3.0) | |
| Diabetic retinopathy (DR) | 3.0 (1.5 - 4.0) | |
| Year of protocol publication | | |
| 2000 or earlier | 1.0 (1.0 – 3.0) | 0.1635 |
| 2001 to 2005 | 2.0 (2.0 - 4.0) | |
| 2006 to 2010 | 3.0 (2.0 - 4.0) | |
| 2011 or later | 2.0 (2.0 - 3.0) | |
| Type of outcome domain specified | | |
| Outcomes specified as primary | 4.0 (3.0 - 4.0) | 0.0001 |
| Outcomes specified as non-primary | 2.0 (1.0 – 3.0) | |
| Type of outcome not specified | 1.0 (1.0 – 2.0) | |

[1] 54 protocols and 3 completed reviews; Median 3.0 (2.0 – 4.0) for outcomes in the 54 protocols and 1.5 (1.0 – 2.0) for outcomes in the 3 reviews (p=0.0627); One protocol did not include any of the outcome domains selected for detailed data extraction

Table 3-4: Examples of incomplete outcome pre-specification

| **Exact text from methods section of protocol** | **Number of completely-specified elements (out of five possible)** |
|---|---|
| *The primary outcome for the review will be visual acuity.* | 1 |
| *When available quality of life data will be described for those with operated and unoperated cataract.* | 1 |
| *Postoperative visual acuity* | 1 |
| *Quality of life* | 1 |
| *Contrast sensitivity* | 1 |
| *Vision-related quality of life at one year* | 2 |
| *Mean IOP* | 2 |

Table 3-5: Frequency of categories of specific metric (element 3) and method of aggregation (element 4) across instances of usage of outcome domains by condition

| Condition/Outcome domain (Number of protocols/Number of instances) | Categories of specific metric (element 3) (% of instances) | | | | Categories of method of aggregation (element 4) (% of instances) | | | |
|---|---|---|---|---|---|---|---|---|
| | Value at a time-point | Time-to-event | Change from baseline | Unclear | Mean | Percent/ proportion | Absolute number | Unclear |
| Glaucoma (22 protocols) | | | | | | | | |
| Quality-of-life (16 instances) | - | - | - | 100 | - | - | - | 100 |
| Visual acuity (13 instances) | 31 | - | 31 | 46 | 8 | 39 | - | 54 |
| Intraocular pressure (20 instances) | 55 | 10 | 25 | 15 | 50 | 10 | - | 40 |
| Disease progression (2 instances) | - | 50 | 50 | - | - | - | - | 100 |
| Cataract (16 protocols) | | | | | | | | |
| Quality-of-life (12 instances) | 17 | - | - | 83 | - | 8 | - | 92 |
| Visual acuity (15 instances) | 53 | - | 20 | 33 | - | 33 | - | 67 |
| Intraocular pressure (1 instance) | 100 | - | - | - | 100 | 100 | - | - |
| Disease progression (3 instances) | - | - | 33 | 67 | - | - | - | 100 |
| Contrast sensitivity (4 instances) | - | - | - | 100 | - | - | - | 100 |
| Age-related macular degeneration (15 protocols) | | | | | | | | |
| Quality-of-life (15 instances) | - | - | - | 100 | - | - | - | 100 |
| Visual acuity (15 instances) | 47 | - | 33 | 40 | 20 | 13 | 13 | 60 |
| Disease progression (6 instances) | - | - | - | 100 | - | - | 17 | 83 |
| Contrast sensitivity (11 instances) | 18 | - | 9 | 73 | 18 | 18 | - | 82 |
| Diabetic retinopathy (4 protocols) | | | | | | | | |
| Quality-of-life (4 instances) | - | - | - | 100 | - | - | - | 100 |
| Visual acuity (4 instances) | 100 | - | 25 | - | - | - | - | 100 |
| Disease progression (4 instances) | 50 | 25 | - | 25 | 25 | - | - | 75 |

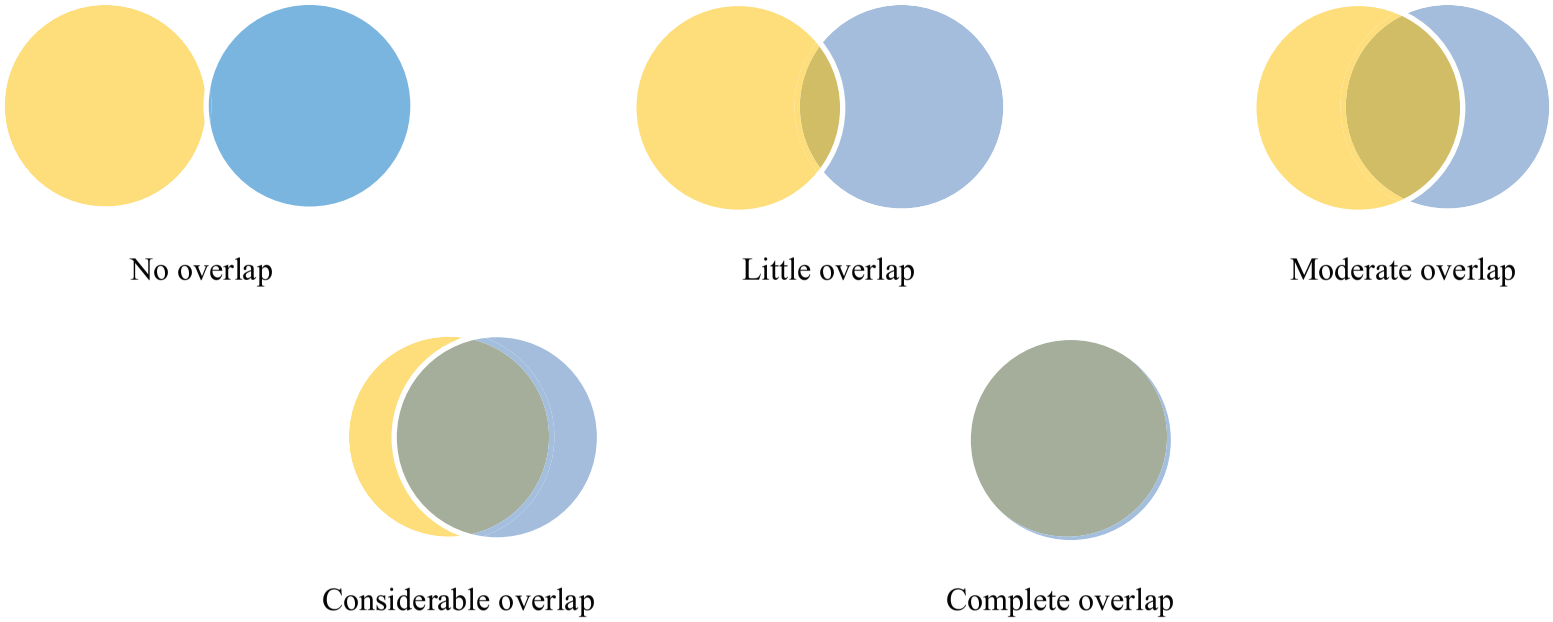
[1] If there was no instance of usage of a certain outcome domain across all reviews addressing a given condition, the above table does not include a row for that outcome domain for that condition.

[2] Percentages are row percentages. Percentages sometimes total more than 100% for an element because some protocols used more than one category for that element.

[3] No reviews used "median" as a method of aggregation.

## FIGURES

Figure 3-1 - Examples of extent of overlap of possible outcome domains chosen by clinical trialists and systematic reviewers




**Legend**

Yellow - Outcomes chosen by clinical trialists
Blue - Outcomes chosen by systematic reviewers
Grey - Outcomes chosen by BOTH clinical trialists and systematic reviewers

Figure 3-2 - Five elements of a completely specified outcome, with anxiety as an example

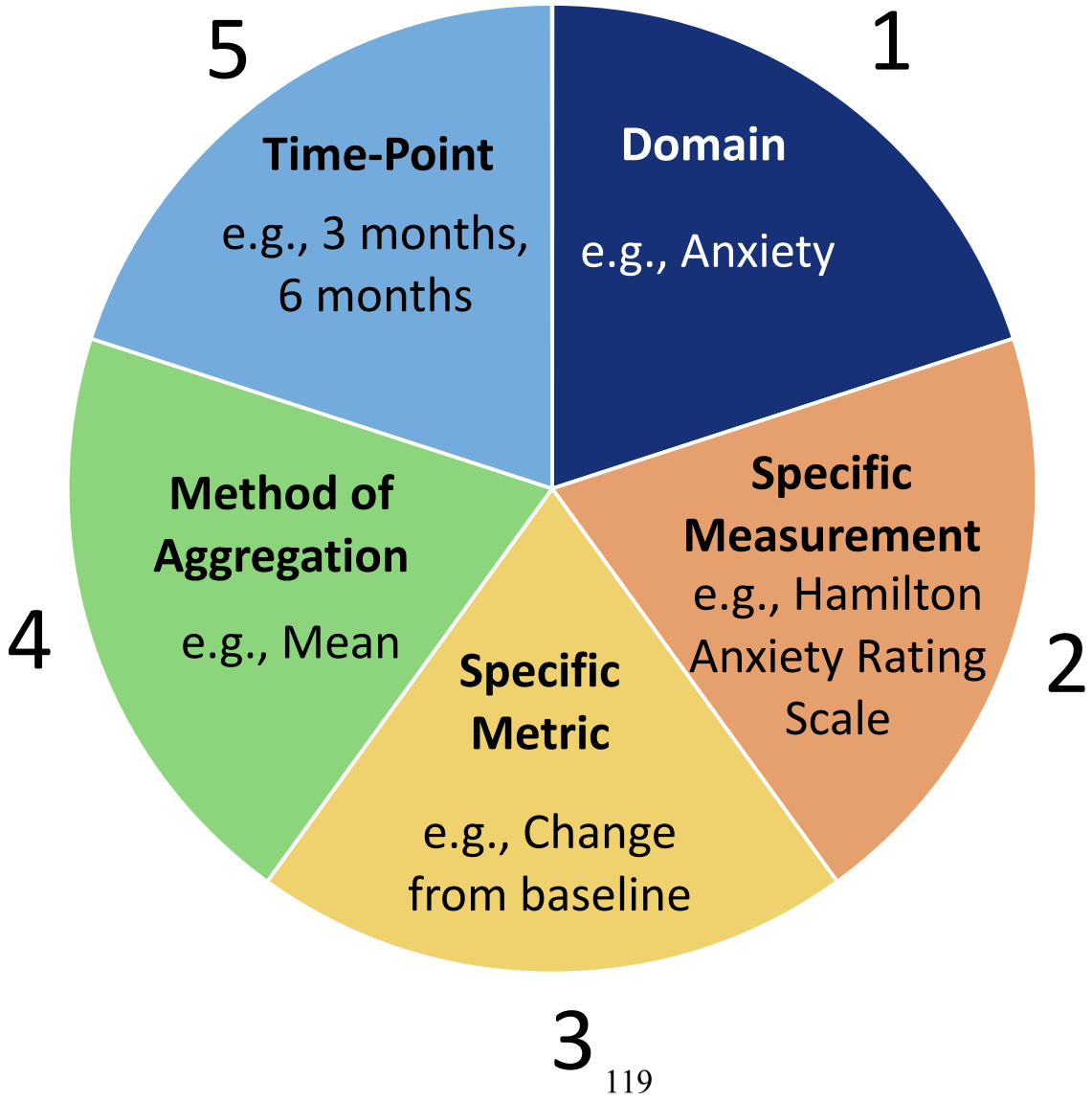
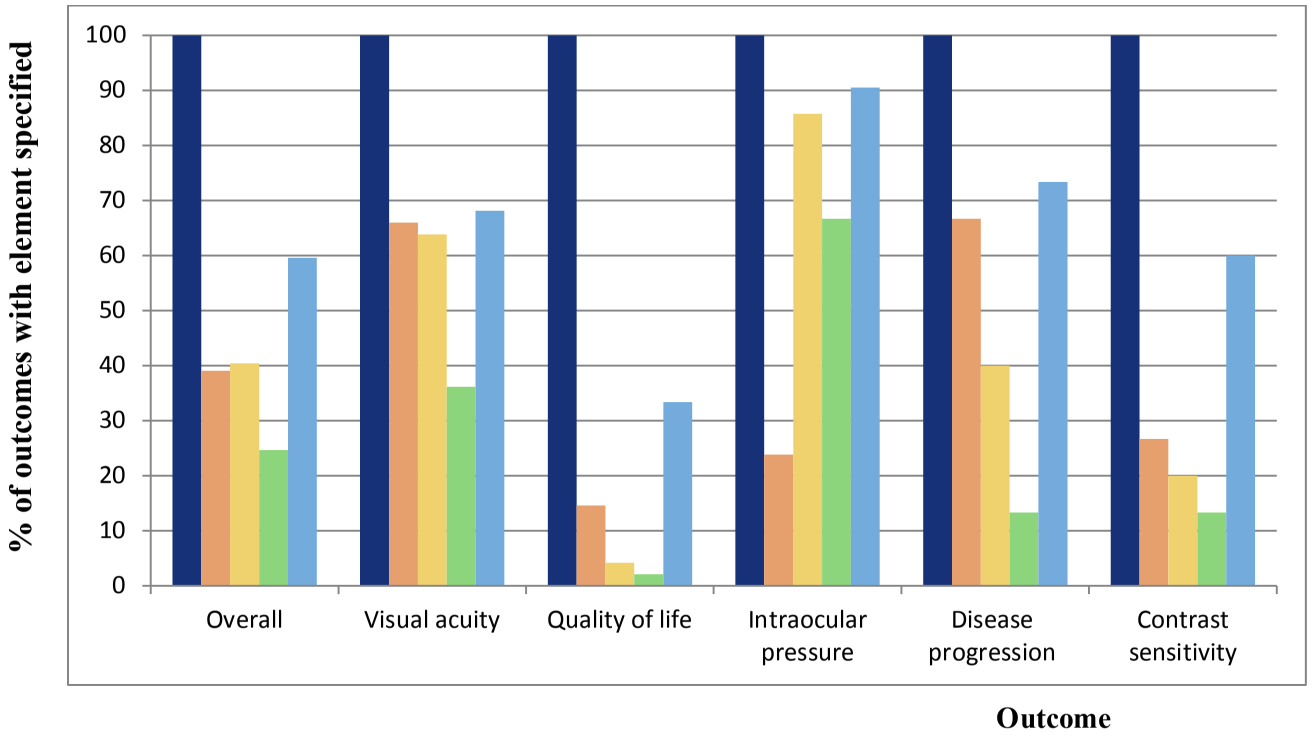


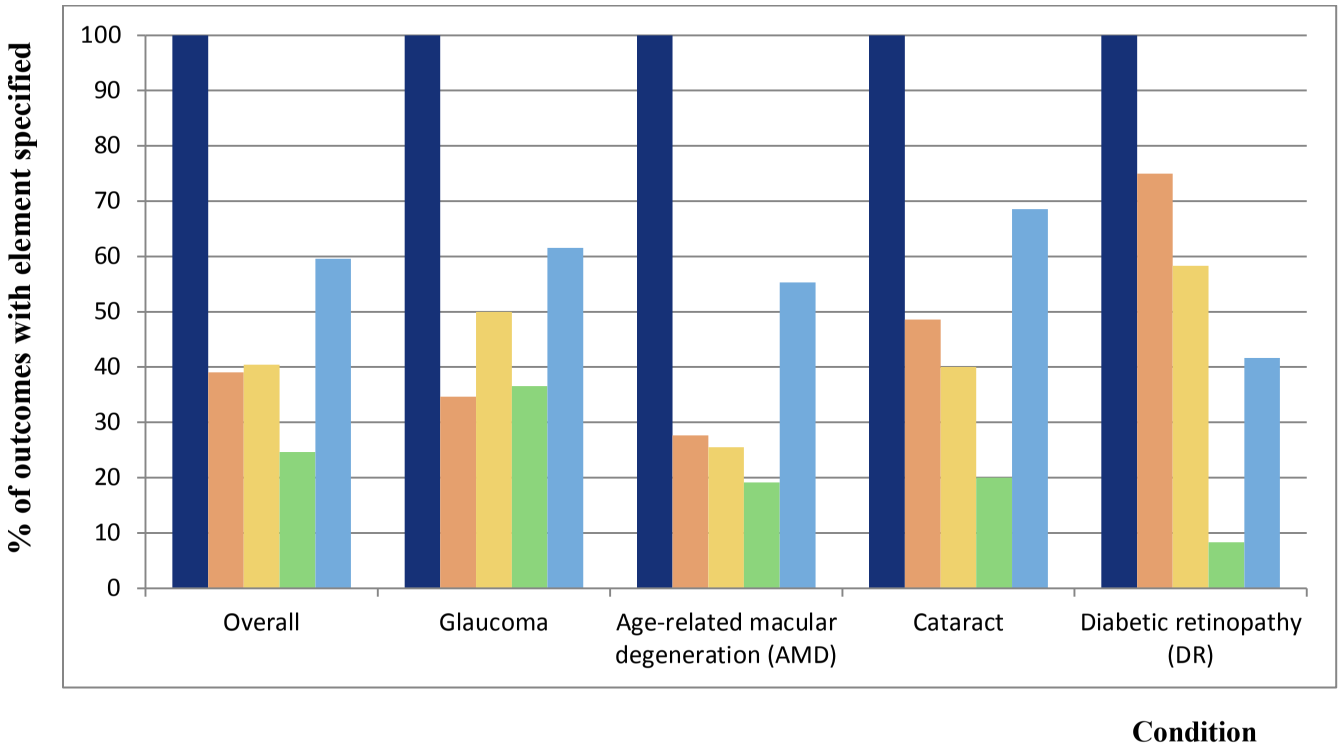
Figure 3-3 - Completeness of specification of outcome elements, by outcome




**Legend**

Navy blue - Domain
Orange – Specific measurement
Yellow – Specific metric
Green – Method of aggregation
Blue – Time-point(s)

Figure 3-4 - Completeness of specification of outcome elements, by condition




**Legend**

Navy blue - Domain
Orange – Specific measurement
Yellow – Specific metric
Green – Method of aggregation
Sky blue – Time-point(s)

# CHAPTER FOUR

## MANUSCRIPT THREE

## Dependability of outcome results in conference abstracts of randomized controlled trials in ophthalmology, and author financial conflicts of interest as a factor associated with full publication

**Dependability of outcome results in conference abstracts of randomized controlled trials in ophthalmology, and author financial conflicts of interest as a factor associated with full publication**

Ian J Saldanha,[1] Roberta W. Scherer, [1] Isabel Rodriguez-Baraquer,[1] Henry D. Jampel,[2] Kay Dickersin[1]

[1] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

[2] Department of Ophthalmology, Johns Hopkins University School of Medicine, Baltimore, MD

## ABSTRACT

### Introduction

Previous research has shown that discrepancies exist between information presented in conference abstracts ("abstracts") and corresponding full publications of the same randomized controlled trial (RCT). Conflicts of interest are prevalent in biomedical research. However, the association between author conflicts of interest and full publication of RCTs is unclear.

### Objectives

Using RCTs in ophthalmology as an example, evaluate: (1) the agreement in reported main outcome results comparing abstracts and their corresponding full publications; and (2) the association between author conflicts of interest and full publication of results presented in abstracts.

### Methods

We included all abstracts presented at the 2001-2004 Association for Research in Vision and Ophthalmology (ARVO) conferences describing results of RCTs. Through electronic searching and emailing abstract authors, we identified the earliest full publication (journal article) containing results of each abstract's main outcome, through 2013. We defined the "main" outcome as the specified primary outcome, or when it was not specified or when more than one were specified, we selected (in this order) the outcome mentioned in the Title, the Objective, or the

outcome mentioned first in the Results. To evaluate agreement between the abstract and its paired full publication in main outcome results, we categorized discordance into qualitative (difference in direction of statistical significance) or quantitative (<10%, 10-20%, >20%, or amount unclear). We used the ARVO classification of conflicts of interest: having financial interest; being an employee of a business with interest; being a consultant to a business with interest; being an inventor/developer with patent; and receiving at least one gift in the past year. Using log-binomials, we calculated the relative risk (RR) of publication associated with each conflict of interest in an overall analysis (model 1) and an interaction analysis (model 2) separately for abstracts with main outcomes statistically significant, not statistically significant, and not reported.

**Results**

We included 513 abstracts, of which 230 (44.8%) were published in full. Median time from presentation to publication was 18 months (interquartile range [IQR]=11 to 33 months). Among the 86 abstract/full publication pairs with the same main outcome at the same time-point, there was either quantitative or qualitative discordance in 47 pairs (54.7%): qualitative discordance in 7 pairs and quantitative discordance in 40 pairs (quantitative discordance <10%, 10%-20%, >20%, and of unclear amount in 14, 5, 14, and 7 pairs respectively). First author reporting at least one conflict of interest was associated with greater likelihood of publication (RR=1.31; 95% CI=1.04 to 1.64) and shorter time to publication (log

rank p=0.026). Specific first author conflicts of interest associated with publication were receiving financial support (RR=1.50; 95% CI=1.19 to 1.90) and receiving at least one gift from industry in the past year (RR=1.42; 95% CI=1.05 to 1.92). These associations remained, irrespective of statistical significance of main outcome results.

**Conclusions**

There was either qualitative discordance or some amount of quantitative discordance in main outcome results in the abstract and the full publication for more than half of the pairs. Irrespective of main outcome results, a declared conflict of interest of the abstract's first author (specifically, receiving financial support and at least one gift from industry in the past year) was associated with full publication of abstracts of RCTs.

## BACKGROUND

There are three main groups of individuals impacted by results of randomized controlled trials (RCTs) presented at conferences: clinicians, patients, and those synthesizing all available research, systematic reviewers. RCT results presented at conferences directly impact clinicians and patients because clinical decisions are sometimes based solely on results presented as conference abstracts ("abstracts").[1-3] For example, Gross and colleagues documented that pre-publication dissemination of results promptly led to substantial changes in clinical practice associated with carotid endarterectomy.[2] RCT results presented at conferences indirectly impact clinicians and patients through systematic reviews. Systematic reviews of RCTs are considered the strongest form of evidence for the effectiveness of healthcare interventions,[4] and underpin evidence-based clinical practice guidelines and other policy-related decision-making. When study results presented in a conference abstract have not been published in full, it is recommended that systematic reviewers include those results.[5-7]

Due to the impact of conference abstracts on clinical decision-making and systematic reviews, it is vital that clinicians, patients, systematic reviewers, and other users of systematic reviews can depend on information presented in abstracts. The concerns are that abstracts do not typically undergo full peer review; might contain preliminary results of RCTs; and might not contain sufficient information to assess the methodological quality of the RCT.[8] Indeed, research has shown that 40% to 62% of RCTs either changed, introduced, or

omitted at least one primary outcome between the protocol and publication stages.[9, 10] Research has also demonstrated discrepancies in analysis methods when comparing an RCT's publication to other documents such as the protocol and trial registry entries. These discrepancies include types of subgroup analyses (61% to 100%) and statistical adjustment (46% to 82%).[10] Similarly, study results have been shown to be discrepant when comparing abstracts with full publications of the same RCT. In ophthalmology, 34% of primary outcomes reported in abstracts had not been specified as such in ClinicalTrials.gov, a major clinical trials registry.[11] Studies in the fields of orthopedics,[12, 13] cardiology,[14] pediatrics,[15] pediatric surgery,[1] and infectious disease[16] have shown that 40% to 60% of RCTs report discrepant results in primary outcome results comparing abstracts and full publications. These comparisons, based on abstracts that reach full publication, shed light on whether abstracts that do not reach full publication should be considered sources of dependable information about RCTs.

Abstracts that do get published are not a random sample of all abstracts, however. We do know that certain study-level characteristics are associated with full publication. Industry funding has been shown to be associated with greater likelihood of full publication [17-19] and, when RCTs are fully published, more favorable study results.[20-23] The impact of study investigator conflicts of interest is less clear, however. Conflicts of interest refer to the set of conditions in which professional judgment concerning a primary interest (such as a patient's

welfare or the validity of research) might be unduly influenced by a secondary interest (such as financial gain).[24, 25]

Real and perceived financial conflicts of interest are prevalent in biomedical research. A survey of academic investigators found that approximately 28% received industry support for their research, 43% received research-related gifts, and 33% had personal financial ties with industry.[26] In a separate survey of clinical practice guideline development panel members, 52% reported having some form of financial association with industry, representing a conflict of interest.[27]

Among RCTs that are published, the presence of conflicts of interest is associated with greater likelihood of the authors' conclusions favoring experimental interventions.[28] The potential impact of conflicts of interest on the likelihood of publication appears complex, though. On one hand, financial gain might threaten impartial judgment,[29] leading to RCT authors selectively publishing RCTs, or rushing them to publication, based on the results (*publication bias)*,[30-32] or even selectively reporting certain outcomes in the publications (*outcome reporting bias*).[9, 10, 32-35] On the other hand, the monetary gain associated with having a financial relationship with industry might facilitate publication, irrespective of study results.

One step of the RCTs results dissemination process during which publication bias, outcome reporting bias, and the impact of conflicts of interest might operate is between presentation of results at a conference and full

publication of those results in journal articles.[36, 37] Conferences offer an important platform for presenting results of RCTs. Presentation is a goal either because it represents an opportunity to disseminate results, or, because the authors wish to attend the conference and the submitted abstract is a means to attend. In most cases, it is difficult to interpret whether publishing in full was a goal of the authors. Indeed, publication not being a goal is frequently reported by authors as a reason for not publishing abstracts in full.[37]

## OBJECTIVES

We conducted a study of RCTs in ophthalmology to evaluate: (1) the agreement in reported main outcome results comparing abstracts and their corresponding full publications; and (2) the association between author conflicts of interest and full publication of results presented in abstracts.

## METHODS

*Included abstracts*

We included all abstracts describing results of RCTs presented at the 2001-2004 Association for Research in Vision and Ophthalmology (ARVO) conferences. This annual conference is the largest international research conference in vision science. Researchers present studies of various designs addressing basic science and various clinical conditions. We considered eligible for this study RCTs addressing any type of intervention for any clinical condition

or for healthy volunteers. We excluded abstracts of non-RCTs or abstracts of RCTs where only non-randomized comparisons were made. Because we did not obtain individual human subjects data, the Johns Hopkins Bloomberg School of Public Health Institutional Review Board (IRB) determined that IRB approval was not needed (IRB #00001810).

*Identifying full publications*

We searched for the earliest full publication (journal article) containing results of each abstract's main outcome. We employed two strategies to identify full publications:

1. We searched the following electronic databases: MEDLINE, Cochrane Central Register of Controlled Trials (CENTRAL), EMBASE, Latin American and Caribbean Health Sciences Literature (LILACS), Web of Science, and SCOPUS. We searched by abstract author (in this order: first, second, last, and then remaining authors), combined with at least one unique search term (derived from the abstract). If no article was identified, we repeated the search with at least two other search terms coupled with each author, in the order described above. We limited the search to two years before the abstract was presented, through June 2013.
2. When full publications were not identified through electronic searches, or if insufficient information was provided in the abstract and/or full publication to confirm a match, we contacted authors of the abstracts via e-mail in April

2009, asking them if the RCT presented in the abstract was published in full. E-mail addresses were retrieved from publicly available sources (e.g., PubMed, Google®, academic websites) and from the ARVO member directory.

*Data extraction*

We developed a data extraction form for both abstracts and full publications using the Systematic Review Data Repository (SRDR) (http://srdr.ahrq.gov). SRDR is an open-access web-based data system that allows structured and organized data extraction through remote and simultaneous access by multiple users. We pilot-tested the data extraction form with both abstracts and full publications, and revised it as needed before data extraction. For each abstract and full publication, we extracted information on author characteristics (authorship order, primary affiliations, and conflicts of interest), study characteristics, study design, participants, interventions, comparisons, and all outcomes. Two investigators extracted data about each abstract and full publication independently, and resolved discrepancies through discussion.

*Identifying the "main" outcome*

Because abstracts and full publications of RCTs do not always specify the "primary" outcome(s), we developed an algorithm *a priori* to classify a reported outcome as the *main* outcome in this order:

1. If only one primary outcome was specified, we selected that outcome as the main outcome.
2. If more than one primary outcome was specified, we selected the first outcome among them with results in the Results as the main outcome.
3. If no primary outcome was specified, we selected the outcome mentioned in the Title or Objective as the main outcome.
4. If no primary outcome was specified and no outcome was mentioned in the Title or Objective, we selected the first outcome with results in the Results section as the main outcome.

*Classifying statistical significance of results for the main outcome*

For the main outcome in each abstract and full publication, we extracted all reported results data at the last available time-point, including individual treatment group ("arm") data, between-arm effect estimate, 95% confidence intervals, and p-value. We considered results for the main outcome as *statistically significant* if the effect estimate, 95% confidence interval, or p-value was statistically significant at the 5% level for at least one reported between-arm comparison for the main outcome (favoring either the experimental arm or the control arm); and *not statistically significant* if not statistically significant at the 5% level for each reported between-arm comparison for the main outcome. When insufficient data were reported for between-arm comparisons, or when the authors only stated that results were statistically significant without reporting the effect

estimate, 95% confidence interval, or p-value, we considered statistical significance of results as *not reported*.

While our main analysis focused on the main outcome, we also evaluated statistical significance for all other outcomes.

*Evaluating agreement in main outcome results between abstracts and corresponding full publications*

For each abstract/full publication pair, we first determined whether the main outcome was the same and reported at the same time-point in both documents. For each pair in which the main outcome was the same and statistical significance of results for the main outcome was reported, we evaluated whether the results agreed.

We defined discordance as either qualitative or quantitative. *Qualitative discordance* refers to a difference in the direction of the effect estimate or statistical significance of the p-value. For example, a difference in direction may imply that the effect estimate/p-value was statistically significant ($p<0.05$) in the abstract and not statistically significant in the full publication (or vice versa), or that one intervention arm was statistically significantly favored in the abstract and another arm was statistically significantly favored in the full publication.

*Quantitative discordance* of an effect estimate for the main outcome refers to any difference in the magnitude but not in the direction of effect estimates, when comparing the abstract with the full publication. Specifically, for each

abstract/full publication pair, we calculated percent difference in reported effect estimates as follows:

$$Percent\ Difference = \left|\frac{Effect\ estimate\ reported\ in\ abstract\ -\ Effect\ estimate\ reported\ in\ full\ publication}{Effect\ estimate\ reported\ in\ full\ publication}\right| X\ 100$$

We categorized quantitative discordance based on the percent differences: <10%, 10%-20%, and >20%. When effect estimates were reported in both the abstract and full publication, we used the reported effect estimates to make the comparison. When effect estimates were not reported in one or both of the two reports in the matched pair, but sufficient information was reported to calculate effect estimates, we calculated effect estimates. When only p-values were reported (without effect estimates or sufficient information to calculate effect estimates), any difference in p-value without a difference in the direction of statistical significance (at the 5% level) also qualified as quantitative discordance; we categorized these instances as 'quantitative discordance – amount unclear'.

*Classifying author conflicts of interest*

We adopted the financial conflicts of interest classification system used by ARVO in 2001 through 2004. This system classifies conflicts of interest into six types: receiving financial support (F); having personal financial interest (I); being an employee of a business with interest (E); being a consultant to a business with interest (C); being an inventor/developer with patent (P); and receiving at least one gift from a business with interest in the past year (R). Further explanation

about these types of conflicts of interest is provided in Box 4-1. For year 2001, conflicts of interest data were presented only as aggregated information for the entire author team. Starting in 2002, ARVO required disclosure of conflicts of interest separately for each author of an abstract, and we extracted all disclosed conflicts of interest for each author.

*Statistical analysis*

To evaluate the agreement in main outcome results in abstract/full publication pairs, we calculated the frequency with which each form of discordance occurred.

To examine the association between conflicts of interest and full publication, we calculated relative risks (RR) of publication of abstracts using log-binomial models. To examine whether the association between conflicts of interest and full publication depended on the statistical significance of results for the abstract's main outcome, we used both overall models (model 1) and interaction models (model 2). Model 2 contained terms for the interaction between the specific conflict of interest and (1) whether or not statistical significance of results for the main outcome was reported in the abstract, and (2) whether or not results for the main outcome in the abstract was statistically significant.

We plotted Kaplan-Meier curves depicting the cumulative probability of full publication of abstracts over time (in months). We conducted log rank significance tests of differences in probability of publication at any time point

during follow-up.[38] We classified abstracts that were published before or at the time of presentation at the conference as being published within one month of presentation.

All statistical analyses were conducted using STATA® 12.

## RESULTS

*Included abstracts*

A total of 20,721 abstracts were presented at ARVO during years 2001-2004, of which 545 abstracts (2.6%) described results of RCTs (Figure 4-1). We excluded 32 of these 545 abstracts because they only described non-randomized comparisons nested within RCTs. We included the remaining 513 abstracts in this study.

*Specification of main outcome in included abstracts*

Among the 513 abstracts, 49 abstracts (9.6%) specified one primary outcome, 15 abstracts (2.9%) specified more than one primary outcome, and 449 abstracts (87.5%) did not specify any primary outcome. We selected the single specified primary outcome as the main outcome for 49 abstracts (9.6%), the outcome mentioned in the Title or Objective as the main outcome for 318 abstracts (62.0%), and the outcome mentioned first in the Results section as the main outcome for the remaining 146 abstracts (28.5%).

*Statistical significance of results for the main outcome in included abstracts*

Statistical significance of results for the main outcome was not reported for 285/513 abstracts (55.6%) (Table 4-1). When it was reported, results for the main outcome were statistically significant in 117/228 abstracts (51.3%). Most of the abstracts reporting statistically significant results for the main outcome favored the experimental rather than the control arm (102/117 abstracts, 87.2%).

*Evaluating agreement in main outcome results between abstracts and corresponding full publications*

Among all 513 abstracts, 230 (44.8%) were published in full. Median time from conference presentation to full publication was 18 months (interquartile range (IQR)=11 to 33 months, range=1 to 90 months). After approximately 60 months, the cumulative proportion of abstracts that was published reached a plateau (Figure 4-2a).

We could not use all 230 pairs of abstracts and corresponding full publications to evaluate agreement in main outcome results. For 190/230 pairs (83.6%), the main outcome was the same in the two reports. Main outcomes were more likely to be the same if the results of the main outcome in the abstract were statistically significant compared with not statistically significant or not reported (RR=1.8; 95% CI=1.09 to 1.28). Among the 190 pairs in which the main outcome was the same, the last available time-point in the abstract was not reported in the full publication for 104 pairs (54.7%). Among these 104 pairs, we determined that the

RCT was ongoing (still following participants) when the abstract was written for 19 pairs (18.3%). For evaluation of agreement in results, we used the 86 pairs for which results for the main outcome were reported in both the abstract and the full publication at the same time-point.

Among the 86 pairs, 39 (45.3%) exactly agreed, while there was some form of discordance in the remaining 47 (54.7%): quantitative discordance in 40 pairs and qualitative discordance in 7 pairs (Figure 4-3, left). Among the 40 pairs with quantitative discordance, discordance was <10% in 14 pairs, 10%-20% in 5 pairs, and >20% in 14 pairs. In 7 pairs, there was quantitative discordance, but the amount was unclear because only p-values were reported in the abstract and/or full publication.

If we defined agreement as exact agreement or <10% of discordance, agreement would be observed for 53 pairs (61.6%), and some form of discordance would be observed for 33 pairs (38.4%) (Figure 4-3, right).

*Association between RCT characteristics and full publication*

Abstracts that were presented as oral presentations were more likely to be published than abstracts presented as posters (RR=1.25; 95% CI=1.01-1.56). For the time period we examined, 2001-2004, ARVO did not require that authors of conference abstracts disclose the study's funding source. A little more than half of the abstracts reported a funding source (271/513 abstracts, 53.0%), and among them, 114/271 abstracts (41.9%) explicitly reported receiving no funding (Table 4-

1). Proportions of abstracts reporting industry, government, and other funding were similar (20.7%, 26.1%, and 21.7% respectively). Most abstracts (361/513 abstracts, 70.4%) did not report whether the RCT was single center or multicenter. Among the abstracts that did, more abstracts described multicenter RCTs (106/152 abstracts, 69.7%) than single center RCTs (46/152 abstracts, 30.3%).

Funding source was associated with full publication. Abstracts describing RCTs that were not funded were less likely to be published compared with other abstracts (RR=0.76; 95% CI=0.58 to 0.99). Abstracts describing RCTs that received industry funding were more likely to be published than those that did not (RR=1.42; 95% CI=1.12 to 1.79). Similarly, abstracts describing RCTs that received government funding were more likely to be published than those that did not (RR=1.31; 95% CI=1.04 to 1.65). These findings are reflected in shorter time to publication for abstracts of funded RCTs compared with abstracts those of unfunded RCTs and abstracts with funding not reported (log rank p=0.021) (Figure 4-2b).

Abstracts describing multicenter RCTs were more likely than single-center RCTs to be published (RR=1.79; 95% CI=1.15 to 2.80). Abstracts of multicenter RCTs also had shorter time to publication than abstracts of single-center RCTs and abstracts where number of centers was not reported (log rank p=0.006) (Figure 4-2c).

*Association between statistical significance of results and full publication*

Abstracts reporting whether results for the main outcome were statistically significant were more likely to be published in full (RR=1.44; 95% CI=1.19 to 1.74), and were published sooner (log rank p<0.001) (Figure 4-2d) compared with abstracts which did not report statistical significance. Among abstracts reporting statistical significance of results for the main outcome (n=228 abstracts), there was no statistically significant association between statistical significance of results and full publication (RR of publication comparing statistically significant with statistically non-significant results=0.97; 95% CI=0.76 to 1.24). Time to publication was similar comparing abstracts with statistically significant and statistically non-significant results for the main outcome (Figure 4-2e).

Abstracts reporting whether results for *any* (main or non-main) outcome were statistically significant were more likely to be published (RR=1.51; 95% CI=1.24 to 1.84) compared with abstracts not reporting statistical significance for any outcome. Among abstracts reporting statistical significance of results for at least one outcome (n=249 abstracts), the likelihood of publication was similar comparing abstracts reporting at least one statistically significant outcome and abstracts reporting no statistically significant outcome (RR=1.06; 95% CI=0.84-1.35).

*Assumption-based analyses of association between statistical significance of results for the main outcome and full publication*

Because a large proportion of the abstracts did not report statistical significance of results for the main outcome, we evaluated the association between statistical significance of results for the main outcome and full publication under five different hypothetical scenarios (Figure 4-4). When we assumed the results for the main outcome were statistically significant in 100% (assumption 1), 75% (assumption 2), or 50% (assumption 3) of abstracts not reporting statistical significance of the main outcome, the associations between significance of results and full publication were statistically significant or borderline statistically significant (RR=1.31, 95% CI=1.07 to 1.60; RR=1.28, 95% CI=1.06-1.55; and RR=1.22; 95% CI=1.00 to 1.48 respectively). When we assumed the main outcome results were statistically significant in 25% of those abstracts (assumption 4), the association between significance of results and full publication was not statistically significant. An inverse association was observed when we assumed the main outcome results were statistically significant in none of those abstracts (RR=0.80; 95% CI=0.65 to 0.99) (assumption 5).

*Affiliations of authors of included abstracts*

The most frequent primary affiliation of abstract first authors was an academic institution (311/505 abstracts, 61.6%) (Table 4-1). Compared with last

authors, first authors less frequently reported primary affiliation with industry (6.1% vs. 17.8%).

*Difference in authorship comparing abstracts and full publications*

Among the 230 pairs of abstracts and corresponding full publications, there was at least one difference in authorship in 202 pairs (87.8%) (Table 4-2). Only middle (i.e., non-first and non-last) authors changed in 44 pairs (19.1%). In 85 pairs, either the first author (33 pairs, 14.4%) or last author (52 pairs, 22.6%) changed, but not both. In 36 pairs, some form of exchange between first and last author took place: they swapped positions in 6 pairs (2.6%), last author became first author in 15 pairs (6.5%), and first author became last author in 15 pairs (6.5%). In 37 pairs (16.7%), neither did the first and last author retain their position nor did they take each other's position.

Overall, the first author and was removed from the full publication in 36/230 pairs (15.6%). The second author was removed in 41/230 pairs (17.8%).

*Conflicts of interest of authors of included abstracts*

For approximately one-third of the abstracts, at least one author reported having at least one conflict of interest (177/513 abstracts, 34.5%). Financial support (99/513 abstracts, 19.3%) and employment by a business with interest (76/513 abstracts, 14.8%) were reported most frequently. Examining the conflicts of interest separately for first and last authors suggests similar distributions to

those of any author. However, compared with last authors, first authors more often reported receiving financial support (18.8% vs. 9.0%), and less often reported employment by a business with interest (4.3% vs. 13.9%).

*Associations between conflicts of interest and full publication – Overall analysis (model 1)*

Overall, when looking at the first author, but not the last author or 'any' author, reporting at least one conflict of interest appeared to be associated with full publication. In model 1, abstracts whose first author reported having at least one conflict of interest had greater likelihood of publication (RR=1.31; 95% CI=1.04 to 1.64) (Table 4-3) and a statistically significantly shorter time to publication (log rank p=0.026) (Figure 4-2f) compared with abstracts whose first author reported no conflicts of interest. A positive association between conflicts of interest and full publication was not observed for the last author or for 'any' author (Table 4-3 and Figures 4-2g and 4-2h).

Regarding specific conflicts of interest, abstracts with first author reporting receiving financial support from industry (RR=1.50; 95% CI=1.19 to 1.90) or at least one gift from industry within the past year (RR=1.42; 95% CI=1.05 to 1.92) were more likely to be published in full compared with abstracts with first author not reporting these conflicts of interest.

*Associations between conflicts of interest and full publication – Interaction analysis (model 2)*

Table 4-3 shows results from an analysis of interaction between conflicts of interest and statistical significance of results for the main outcome in the abstract (model 2). Generally, model 2 estimates of RR of publication for each subgroup of abstracts were similar to model 1. This was supported by the fact that none of the interaction terms were statistically significant, suggesting that the associations between conflict of interest and publication did not differ based on whether the statistical significance of results for the main outcome in the abstract were not statistically significant, statistically significant, or not reported. For example, the first author receiving financial support from industry was associated with 50% higher likelihood of publication in model 1, while the corresponding estimates were 44%, 37%, 11% for abstracts with statistically significant results, without statistically significant results, and statistical significance of results not reported, respectively.

## DISCUSSION

In this longitudinal study of 513 abstracts of RCTs presented at the world's largest ophthalmology research conference (ARVO), we have demonstrated that there was either qualitative or quantitative discordance in results for the main outcome comparing the abstract with the full publication for approximately 55% of pairs. As regards conflicts of interest, abstracts with first authors reporting at

least one conflict of interest were 31% more likely to be published, and were published sooner, compared with abstracts with first author not reporting any conflict of interest. This association did not depend on whether the results for the main outcome in the abstract were statistically significant, statistically non-significant, or whether the statistical significance of the association was reported.

*Dependability of conference abstarcts: Implications for clinicians, patients, and systematic reviewers*

The discordance between results of main outcomes presented in abstracts and full publications is worrying. For almost one in five (16.4%) pairs of abstracts and full publications, the main outcomes were different in the two reports. When the main outcome was indeed the same, results for the outcome's last time-point in the abstract were not reported in the full publication for more than half of the pairs (54.7%). When the same main outcome was reported at the same time-point, approximately 8% of abstract/full publication pairs reported qualitatively different results in the two reports. This implies that if a decision-maker was using an RCT to make a treatment decision, approximately one in twelve such treatment decisions based on the same outcome could differ depending on whether the decision-maker looked at the abstract or the full publication.

Even when there was no qualitative discordance, almost half of the abstract/full publication pairs reported quantitatively different results for the same main outcome. When discordance could be quantified, we classified it into three

categories based on cutoffs at 10% and 20%. We recognize that these cutoffs are arbitrary, but the amount of discordance that might be tolerated likely depends on the field and the context in which the results are used. For example, estimates of RR of 1.3 and 1.5 from a single RCT might not readily be considered meaningful by decision-makers, but might impact whether the summary effect estimate might or might not overlap null in a meta-analysis. This might especially be true if effect estimates obtained from abstracts are based on preliminary analyses, and therefore, might be based on smaller sample sizes and have wider confidence intervals. In our study, results being preliminary could have explained for less than 20% of the discrepancies we observed. Discrepancies become especially problematic for a new clinical intervention, where, for example, a health plan may ask systematic reviewers for a summary of available evidence so as to allow a treatment reimbursement decision. In this example, if most of the evidence is reported in abstracts, and the individual effect estimates are consistently inflated, this could impact decision-making.

Our findings of discordances between abstracts and full publications (among published abstracts) suggest that it is conceivable that the dependability of abstracts might be even less for abstracts that are not published.

The Institute of Medicine (IOM), the Cochrane Collaboration, and the Agency for Healthcare Research and Quality (AHRQ) all recommend that, when study results are not available from publications, systematic reviews include results from both unpublished sources, such as abstracts.[5-7] In a systematic review, there are

a few possible scenarios and typical strategies for a certain outcome from a certain RCT:

- Scenario 1: Results for the outcome are only available from an abstract. The systematic reviewer typically includes the data from the abstract in the meta-analysis, and should conduct a sensitivity analysis to examine whether excluding the data affects the meta-analytic effect estimate.[8]
- Scenario 2: Results for the outcome are available from both an abstract and a full publication, and the data exactly agree. This is the ideal scenario, and it does not matter from where the systematic reviewer obtains the data.
- Scenario 3: Results for the outcome are available from both an abstract and a full publication, and there is quantitative discordance. The systematic reviewer typically includes the data from the full publication. However, in doing so, the systematic reviewer makes the implicit assumption that the full publication includes the more valid data.
- Scenario 4: Results for the outcome are available from both an abstract and a full publication, and there is qualitative discordance. The systematic reviewer typically makes the same assumption as in Scenario 3 and includes data from the full publication.

The findings in our study suggest that when results presented in an abstract are published in full (i.e., when scenarios 2, 3, and 4 are possible), discordance (scenarios 3 and 4) is more common than exact agreement (scenario 2). The existence of discordance has two main implications. First, when only an abstract is

available (scenario 1), the systematic reviewer should be cautious about including the results therein, running the appropriate sensitivity analysis.[8] The documented high prevalence of discordances in our study and previous research also supports this approach.[1, 12-16] Second, when both abstract and full publication are available and the data are discordant (scenarios 3 and 4), the systematic reviewer should contact the authors, seeking clarification about the results before proceeding. This is especially critical in instances of qualitative discordance (scenario 4).

*Conflicts of interest, statistical significance of main outcome results in abstracts, and full publication*

Prior research has demonstrated that outcomes in RCTs are selectively reported [33-35] or under-reported [10] based on the direction of the results, leading to outcome reporting bias. Our finding of no association appears to be at odds with this and other research focusing specifically on publication of the main outcome.[19, 39, 40] We believe that the likely reason why our findings differed was that statistical significance of results for the main outcome was not reported for more than half of the abstracts in our study (55.6%), and we did not consider these abstracts in our main analysis. Thus, the abstracts in our study not reporting the statistical significance of results for the main outcome are likely not a random sub-set of all abstracts in our study. Indeed, when we subjected our data to different assumptions for the missing results, we observed results similar to those

found by others. For example, even if we assumed that 50% of abstracts with statistical significance not reported were, in fact, statistically significant (assumption 3), abstracts with statistically significant results for the main outcome were 22% more likely to be published than abstracts with non-statistically significant results for the main outcome. Given what we know about outcome reporting bias, and that approximately half of all reported main outcomes in our study were statistically significant, it is likely that in abstracts not reporting statistical significance, less than half of the main outcomes were actually statistically significant (therefore, assumptions 1 and 2 are more reasonable).

Our finding that conflicts of interest of the first author impact publication more than those of the last author is intriguing. We explored a few potential reasons why this might have occurred. First, financial support in the form of research funding provided to first authors of abstracts might encourage and enable time spent on developing a full publication, and, theoretically, the first author would devote the most time to writing up a full publication. In our study, compared with last authors, first authors were less likely to report industry as a primary affiliation (6.1% vs. 17.8%) and employment by a business with interest as a conflict of interest (4.3% vs. 13.9%). Yet compared with last authors, first authors, also reported receiving more financial support (18.8% vs. 9.0%) and gifts from industry in the past year (9.0% vs. 3.8%). The decision to ask an academic, not an industry employee, to serve as first author may be strategic in terms of influence on opinion leaders. Indeed first authors were slightly more frequently

affiliated with academic institutions than last authors (61.6% vs. 54.1%) in our study. It may be that academics are inclined to serve as abstract authors when they can be first and when publishing in full is an option. Research publications are generally important to advancement in academic careers, which might have motivated first authors affiliated with academic institutions to publish their results in full. Second, it is possible that a greater proportion of first authors than last authors in our study were junior researchers within academic institutions. Submission of abstracts is sometimes seen by junior researchers as means to attend conferences. Although we were unable to examine this directly because ARVO did not collect information from authors about their ranks/positions, we examined this indirectly by calculating the frequency with which first and last authors were removed in the full publications (assuming that removal of the first author would suggest that the first author might have been a junior researcher). The proportion of full publications that removed the first author (36%) was similar to the proportion that dropped the last author (41%), however, suggesting that it is unlikely that first authors were more often junior researchers. Third, it is possible that authors of abstracts in our study were different in some way from authors who have been examined in previous studies examining conflicts of interest. For example, only 33.0% and 28.1% of first and last authors respectively in our study reported at least one conflict of interest. These percentages have been reported to be higher in previous studies among academic investigators [26] and clinical practice guideline developers.[27] Fourth, it is possible that authors in our study

under-reported conflicts of interest. Although conflicts of interest disclosure was required by ARVO, it was self-reported by authors.

Findings from our interaction analysis suggest that, irrespective of statistical significance of results, conflicts of interest of the first author were associated with increased likelihood that authors went beyond presenting RCT results at the conference by publishing those results in full.

*Low rates of publication: Implications for researchers conducting clinical trials*

Even after a long follow-up duration in our study, only 44.5% of abstracts were published in full. Publication is important for a number of reasons. At the very least, non-publication of RCT results is a waste of resources,[41-43] and worse, is arguably scientific misconduct.[19, 44, 45] RCT participants volunteer for clinical trials with the understanding that their participation advances science; non-publication is a violation of that understanding.[19] In addition, publication is important because, as we have demonstrated, there often is discordance between results presented in abstracts and in full publications. Assuming that when discordance exists between results reported in an abstract and a full publication, the full publication contains the more 'valid' results, it is important that researchers writing abstracts be more vigilant. Of course, this assumption does not always hold. In either event, researchers should be vigilant about presenting valid data at conferences. Further, low rates of publication and publication bias imply

that those summarizing the findings of RCTs are only able to access a biased subset of all RCTs.[19]

Even in the current era of increased awareness of the importance of open access to clinical trial data, full publication remains a key form of dissemination of study results. Studies published in full generally go through peer-review, reach a wider audience than abstracts, and contain greater detail and nuance than abstracts. This detail and nuance allows the readers of publications to critically appraise the results and interpret them in context.[8] Researchers must recognize the importance of publishing their results and overcome the barriers of lack of time and low priority.[17, 19, 37, 46, 47] Novel incentive structures for publication that go beyond monetary benefit and career advancement should be developed.

*Strengths of our study*

Our study has a number of strengths. First, we examined all abstracts of RCTs presented at the world's largest research conference in ophthalmology (ARVO) over four consecutive years. To their credit, the organizers of ARVO required conflicts of interest disclosure by authors of all submitted abstracts. This allowed us a large sample size of 513 abstracts. For 400 abstracts (years 2002-2004), conflicts of interest information were available for each author, allowing us to separately examine conflicts of interest for first and last authors. Second, we did not selectively include abstracts based on clinical condition or type of intervention; we included all abstracts, provided participants were randomized.

Third, we had a long follow-up of the abstracts for full publication (maximum follow-up ranging from 110 to 146 months). This comfortably exceeds 48-60 months, which is the time by which the majority of RCTs that would eventually be published are published.[19, 36] Fourth, we employed two strategies to identify full publications: electronic searching of multiple databases and contacting abstract authors via email. It is unlikely that we have missed a substantial number of full publications.

*Limitations to our study*

Our study also has a few potential limitations. First, only 9.6% of abstracts specified an outcome as "primary". However, we *a priori* developed an algorithm to determine the main outcome for those abstracts that did not. Although this algorithm is likely consistent with how most readers of abstracts judge what is meant to be interpreted as the most important outcome, our definition may have influenced our findings related to the "main" outcome. Second, statistical significance of the main outcome was specified for only 228/513 abstracts (44.4%), and only 86/230 abstract/full publication pairs (37.4%) had the same main outcome at the same time-point, and reported statistical significance in both reports. Nevertheless, 86 pairs is a sizeable number for our conclusions. Third, a tradeoff for our study's strength of a long duration of follow-up is that the abstracts we examined were presented at ARVO from 2001-2004, prior to initiatives such as compulsory registration of RCTs. Registration of RCTs at

ClinicalTrials.gov or at registries within the WHO International Clinical Trials Registry Platform (ICTRP) (http://www.who.int/ictrp/en/) could mitigate the impact of publication bias and outcome reporting bias in more recent abstract/full publication pairs. In addition, the CONSORT for Abstracts extension of the CONSORT Statement includes clear specification of the primary outcome as an essential checklist item.[48] This might have improved the reporting of outcomes in more recent abstract/full publication pairs.

*Conclusions*

We have reported qualitative discordance or some amount of quantitative discordance in main outcome results for more than half of the abstract/full publication pairs. Researchers presenting at conferences should be vigilant about the accuracy of the results they present; and clinicians, patients, and systematic reviewers should be aware of these discordances. We have shown that the conflicts of interest most associated with full publication of abstracts are conflicts of interest of the abstract's first author, irrespective of whether the results for the main outcome in the abstract were statistically significant, not statistically significant, or if statistical significance was not reported.

**ACKNOWLEDGEMENTS**

We thank Shilpa Viswanathan, PhD for comments on previous versions of this manuscript. We thank Nicole Fusco, Jack Kent, Shauna Linn, Diana Lock, Ju

Park, Benjamin Rouse, Elizabeth Ssemanda, and Cesar Ugarte-Gil for their assistance with data extraction. This study was funded by the National Eye Institute (grant number 1U01EY020522) and by the Cochrane Eyes and Vision Group US Satellite.

## REFERENCES


1. Weintraub WH. Are published manuscripts representative of the surgical meeting abstracts? An objective appraisal. *J Pediatr Surg* 1987; 22:11–3.

2. Gross CP, Steiner CA, Bass EB, Powe NR. Relation between prepublication release of clinical trial results and the practice of carotid endarterectomy. *JAMA* 2000; 284:2886–93.

3. Falagas ME, Rosmarakis ES. Clinical decision-making based on findings presented in conference abstracts: is it safe for our patients? *Eur Heart J* 2006; 27(17):2038-9.

4. Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ, Green L, Naylor CD, Wilson MC, Richardson WS. Users' Guides to the Medical Literature: XXV. Evidence-based medicine: principles for applying the Users' Guides to patient care. Evidence-Based Medicine Working Group. *JAMA* 2000 Sep 13; 284(10):1290-6.

5. Institute of Medicine. Initial National Priorities for Comparative Effectiveness Research. 2009. Available online at: http://www.iom.edu/~/media/Files/Report%20Files/2009/ComparativeEffectivenessResearchPriorities/CER%20report%20brief%2008-13-09.pdf. Accessed May 13, 2015.

6. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. Available online at: www.cochrane-handbook.org. Accessed May 13, 2015.

7. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov. Accessed May 13, 2015.

8. Dundar Y, Dodd S, Dickson R, Walley T, Haycox A, Williamson PR. Comparison of conference abstracts and presentations with full-text articles in the health technology assessments of rapidly evolving technologies. *Health Technol Assess* 2006; 10(5) :iii-iv, ix-145.

9. Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, et al. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev* 2011*;* Issue 1. Art. No.: MR000031. DOI: 10.1002/14651858.MR000031.pub2.

10. Dwan K, Gamble C, Williamson PR, Kirkham JJ, Reporting Bias Group. Systematic review of the empiral evidence of study publication bias and outcome reporting bias – an updated review. *PLoS One* 2013; 8(7):e66844. doi: 10.1371/journal.pone.0066844.

11. Scherer RW, Huynh L, Ervin A-M, Dickersin K. Using ClinicalTrials.gov to supplement information in ophthalmology conference abstracts about trial outcomes: a comparison study. PLoS One. In press.

12. Bhandari M, Devereaux PJ, Guyatt GH, Cook DJ, Swiontkowski MF, Sprague S, Schemitsch EH. An observational study of orthopaedic abstracts and subsequent full-text publications. *J Bone Joint Surg Am* 2002; 84-A:615–621.

13. Kleweno CP, Bryant WK, Jacir AM, Levine WN, Ahmad CS. Discrepancies and rates of publication in orthopedic sports medicine abstracts. *Am J Sports Med* 2008; 36(10):1875-9.

14. Toma M, McAlister FA, Bialy L, Adams D, Vandermeer B, Armstrong PW. Transition from meeting abstract to full-length journal article for randomized controlled trials. *JAMA* 2006; 295:1281–1287.

15. Klassen TP, Wiebe N, Russell K, Stevens K, Hartling L, Craig WR, Moher D. Abstracts of randomized controlled trials presented at the society for pediatric research meeting: an example of publication bias. *Arch Pediatr Adolesc Med* 2002; 156:474–479.

16. Rosmarakis ES, Soteriades ES, Vergidis PI, Kasiakou SK, Falagas ME. From conference abstract to full paper: differences between data presented in conferences and journals. *FASEB J* 2005; 19:673–680.

17. Dickersin K, Min Y-I, Meinert CL. Factors influencing publication of research results: Follow-up of applications submitted to two institutional review boards. *JAMA* 1992; 267(3):374–8.

18. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ* 1997; 315:640–645.

19. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev* 2007; Issue 2. Art. No.: MR000005. DOI: 10.1002/14651858.MR000005.pub3.

20. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA* 2003; 289(4):454-65.

21. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008; 358(3):252-60.

22. Sismondo S. Pharmaceutical company funding and its consequences: a qualitative systematic review. *Contemporay Clinical Trials* 2008; 29(2):109–13.

23. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2012; Issue 12. Art. No.: MR000033. DOI: 10.1002/14651858.MR000033.pub2.

24. Thompson D. Understanding financial conflicts of interest. *N Engl J Med.*1993; 329:573-576.

25. International Committee of Medical Journal Editors (ICMJE). Author responsibilities-Conflicts of interest. Available at http://www.icmje.org/recommendations/browse/roles-and-responsibilities/author-responsibilities--conflicts-of-interest.html. Accessed May 13, 2015.

26. Campbell EG, Louis KS, Blumenthal D. Looking a gift horse in the mouth: corporate gifts support life sciences research. *JAMA* 1998; 279:995-9.

27. Neuman J, Korenstein D, Ross JS, Keyhani S. Prevalence of financial conflicts of interest among panel members producing clinical practice guidelines in Canada and United States: cross sectional study. *BMJ* 2011. Oct 11;343:d5621. doi: 10.1136/bmj.d5621.

28. Kjaergard LL, Gluud C. Citation bias of hepato-biliary randomized clinical trials. *J Clin Epidemiol* 2002; 55:407-10.

29. Johnston J. Conflict of Interest in Biomedical Research. In *From Birth to Death and Bench to Clinic: The Hastings Center Bioethics Briefing Book for Journalists, Policymakers, and Campaigns*, ed. Mary Crowley (Garrison, NY: The Hastings Center, 2008), 31-34.

30. Simes RJ. Publication bias: The case for an international registry of clinical trials. *J Clin Oncol* 1986; 4:1529–41.

31. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990; 263: 1385-9.

32. Song F, Parekh S, Hooper, L, Loke Y, et al. Dissemination and publication of research findings: an update of related biases. *Health Technol Assess* 2010;14(8).

33. Chan AW, Hrobjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004; 291(20):2457-2465.

34. Chan AW, Altman D. Identifying outcome reporting bias in randomized trials on PubMed: review of publications and survey of authors. *BMJ* 2005; 330(7494):753.

35. Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *N Engl J Med* 2009; 361:1963-1971.

36. von Elm E, Costanza MC, Walder B, Tramèr MR. More insight into the fate of biomedical meeting abstracts: a systematic review. *BMC Medical Res Methodol* 2003; 3(1):12.

37. Scherer RW, Ugarte-Gil C, Schmucker C, Meerpohl JJ. Authors report lack of time as main reason for unpublished research presented at biomedical conferences: a systematic review. *J Clin Epidemiol* 2015; pii: S0895-4356(15)00065-7.

38. Bland JM, Altman DG. The logrank test. *BMJ* 2004; 328(7447): 1073.

39. Ioannidis J. Effect of the statistical significance of results on the time to completion and publication of randomised efficacy trials. *JAMA* 1998; 279:281–6.

40. Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database Syst Rev* 2007; Issue 2. [DOI: 10.1002/14651858.MR000011.pub2]

41. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, Krumholz HM, Ghersi D, van der Worp HB. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014; 383(9913):257-66.

42. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; 374(9683):86-9.

43. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, Salman RA, Chan AW, Glasziou P. Biomedical research: increasing value, reducing waste. *Lancet* 2014; 383(9912):101-4.

44. Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990; 263:1405–8.

45. Antes G, Chalmers I. Under-reporting of clinical trials is unethical. *Lancet* 2003;361:978–9.

46. Weber EJ, Callaham ML, Wears RL, Barton C, Young G. Unpublished research from a medical specialty meeting: Why investigators fail to publish. *JAMA* 1998; 280(3): 257–9.

47. Callaham ML, Wears RL, Weber EJ, Barton C, Young G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *JAMA* 1998; 280(3):254–7.

48. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF and the CONSORT Group (2008) CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008; 371(9609):281-3.

## BOXES

Box 4-1: Definitions for financial conflicts of interest as provided by ARVO

| Conflict of interest | Description |
|---|---|
| Receiving financial support | This category includes funding received through employing institution support or from a not-for-profit or competing company, in the form of research funding/services (e.g. protein sequencing) at no cost, support being the subject matter of the presentation/publication. |
| Personal financial interest | This category includes being an investor in a company or competing company other than through a mutual or retirement fund, which provides a product, service, process, or equipment that is the subject matter of the presentation/publication. |
| Employee of a business with interest | This category includes being an employee of a company or competing company with a business interest that is the subject matter of the presentation/publication. |
| Consultant to a business with interest | This category includes being/having been a consultant for a company or competing company with a business interest that is the subject matter of the presentation/publication. |
| Inventor/developer with patent | This category includes being an inventor/developer designated on a patent, patent application, copyright or trade secret, whether or not the patent, copyright, etc. is presently licensed or otherwise commercialized, which is the subject matter of the presentation/ publication or could be in competition with the technology described. |
| Receiving at least one gift in the past year | This category includes having received at least one gift in kind, honoraria, or travel reimbursement valued at over $1,000 in the last 12 months from a company or competing company which provides a product, service, process, or equipment that is the subject matter of the presentation/publication. |

# TABLES

Table 4-1: RCT characteristics, author characteristics, and main outcome results of abstracts of RCTs presented at ARVO conferences during years 2001-2004, overall and by whether or not the abstract was published in full

| Characteristics | All abstracts n (%**) | Abstracts NOT published in full n (%**) | Abstracts PUBLISHED in full n (%**) |
|---|---|---|---|
| **CHARACTERISTICS OF RCTs** | | | |
| **Presentation at ARVO** | N=513 | N=283 | N=230 |
| Poster | 418 (81.5) | 239 (84.5) | 179 (77.8) |
| Oral | 95 (18.5) | 44 (15.5) | 51 (22.2) |
| **Funding** | N=513 | N=283 | N=230 |
| Not reported | 241 (47.0) | 137 (48.4) | 104 (45.2) |
| Reported | 272 (53.0) | 146 (51.6) | 126 (54.8) |
| At least one funding source | 158 (58.1) | 73 (50.0) | 85 (67.5) |
| Industry (pharmaceutical or other)* | 56 (20.7) | 22 (15.1) | 34 (27.0) |
| Government* | 71 (26.1) | 31 (21.2) | 40 (31.8) |
| Other* | 59 (21.7) | 30 (20.6) | 29 (23.0) |
| No funding | 114 (41.9) | 73 (50.0) | 41 (32.5) |
| **Number of centers** | N=513 | N=283 | N=230 |
| Not reported | 361 (70.4) | 208 (73.5) | 153 (66.5) |
| Reported | 152 (29.6) | 75 (26.5) | 77 (33.5) |
| Single center | 46 (30.3) | 31 (41.3) | 15 (19.5) |
| Multicenter | 106 (69.7) | 44 (58.7) | 62 (80.5) |
| **CHARACTERISTICS OF FIRST AUTHORS** | | | |
| **Primary affiliation** | N=513 | N=283 | N=230 |
| Not reported | 8 (1.6) | 7 (2.5) | 1 (0.4) |
| Reported | 505 (98.4) | 276 (97.5) | 229 (99.6) |
| Academic | 311 (61.6) | 156 (56.5) | 155 (67.7) |
| Industry | 31 (6.1) | 18 (6.5) | 13 (5.7) |
| Hospital/Clinic | 123 (24.4) | 76 (27.5) | 47 (20.5) |
| Other | 40 (7.9) | 26 (9.4) | 14 (6.1) |
| **Conflicts of interest (years 2002-2004 only)** | N=400 | N=229 | N=171 |
| Not reported | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Reported | 400 (100.0) | 229 (100.0) | 171 (100.0) |
| At least one conflict of interest | 132 (33.0) | 65 (28.4) | 67 (39.2) |
| F - Financial support* | 75 (18.8) | 31 (13.5) | 44 (25.7) |
| I - Personal finance interest* | 4 (1.0) | 2 (0.9) | 2 (1.2) |
| E - Employee of business with interest* | 17 (4.3) | 13 (5.7) | 4 (2.3) |
| C - Consultant of business with interest* | 43 (10.8) | 23 (10.0) | 20 (11.7) |
| P - Inventor/developer with patent* | 8 (2.0) | 5 (2.2) | 3 (1.8) |
| R - Received gifts within the past year* | 36 (9.0) | 15 (6.6) | 21 (12.3) |
| No conflict of interest | 268 (67.0) | 164 (71.6) | 104 (60.8) |

| Characteristics | All abstracts | | Abstracts NOT published in full | | Abstracts PUBLISHED in full | |
|---|---|---|---|---|---|---|
| | n (%**) | | n (%**) | | n (%**) | |
| **CHARACTERITICS OF LAST AUTHORS** | | | | | | |
| **Primary affiliation** | N=513 | | N=283 | | N=230 | |
| Not applicable (abstract had one author) | 28 | (5.5) | 8 | (2.8) | 20 | (8.7) |
| Not reported | 8 | (1.6) | 7 | (2.5) | 1 | (0.4) |
| Reported | 477 | (93.0) | 268 | (94.7) | 209 | (90.9) |
| Academic | 258 | (54.1) | 140 | (52.2) | 118 | (56.5) |
| Industry | 85 | (17.8) | 39 | (13.8) | 46 | (22.0) |
| Hospital/Clinic | 106 | (22.2) | 69 | (25.8) | 37 | (17.7) |
| Other | 28 | (5.9) | 20 | (7.1) | 8 | (3.8) |
| **Conflicts of interest (years 2002-2004 only)** | N=400 | | N=229 | | N=171 | |
| Not applicable (abstract had one author) | 21 | (5.3) | 8 | (3.5) | 13 | (7.6) |
| Not reported | 12 | (3.0) | 6 | (2.6) | 6 | (3.5) |
| Reported | 367 | (91.7) | 215 | (93.9) | 152 | (88.9) |
| At least one conflict of interest | 103 | (28.1) | 56 | (26.1) | 39 | (25.7) |
| F - Financial support* | 33 | (9.0) | 21 | (9.8) | 12 | (7.9) |
| I - Personal finance interest* | 1 | (0.3) | 1 | (0.5) | 0 | (0.0) |
| E - Employee of business with interest* | 51 | (13.9) | 28 | (13.0) | 23 | (15.1) |
| C - Consultant of business with interest* | 18 | (4.9) | 8 | (3.7) | 10 | (5.9) |
| P - Inventor/developer with patent* | 5 | (1.4) | 3 | (1.4) | 2 | (1.2) |
| R - Received gifts within the past year* | 14 | (3.8) | 7 | (3.3) | 7 | (4.6) |
| No conflict of interest | 272 | (74.1) | 159 | (74.0) | 113 | (74.3) |
| **CONFLICTS OF INTEREST OF ANY AUTHOR** | | | | | | |
| **Conflicts of interest** | N=513 | | N=283 | | N=230 | |
| Not reported | 0 | (0.0) | 0 | (0.0) | 0 | (0.0) |
| Reported | 513 | (100.0) | 283 | (100.0) | 230 | (100.0) |
| At least one conflict of interest | 177 | (34.5) | 91 | (32.2) | 86 | (37.4) |
| F - Financial support* | 99 | (19.3) | 49 | (17.3) | 50 | (21.7) |
| I - Personal finance interest* | 6 | (1.2) | 4 | (1.4) | 2 | (0.9) |
| E - Employee of business with interest* | 76 | (14.8) | 40 | (14.1) | 36 | (15.7) |
| C - Consultant of business with interest* | 66 | (12.9) | 32 | (11.3) | 34 | (14.8) |
| P - Inventor/developer with patent* | 13 | (2.5) | 8 | (2.8) | 5 | (2.2) |
| R - Received gifts within the past year* | 55 | (10.7) | 23 | (8.1) | 32 | (13.9) |
| No conflict of interest | 336 | (65.5) | 192 | (68.9) | 144 | (62.6) |
| **MAIN OUTCOME RESULTS** | | | | | | |
| **Main outcome - Statistical significance** | N=513 | | N=283 | | N=230 | |
| Not reported | 285 | (55.6) | 178 | (63.9) | 107 | (46.5) |
| Reported | 228 | (44.4) | 105 | (37.1) | 123 | (53.5) |
| Not statistically significant | 111 | (48.7) | 52 | (49.5) | 59 | (48.0) |
| Statistically significant | 117 | (51.3) | 53 | (50.5) | 64 | (52.0) |

* More than one option could apply to each abstract.

** Percentages are column percentages. Percentages in shaded rows are calculated with n reported as the denominator.

Table 4-2: Differences in authorship order comparing 230 conference abstracts with their corresponding full publications

| Status | Number of pairs N (%) |
|---|---|
| No change in authorship | 28 (12.2) |
| Middle authors changed, FIRST and LAST authors remained unchanged | 44 (19.1) |
| LAST author changed, FIRST author remained unchanged | 52 (22.6) |
| LAST author moved to be a middle author | 26 |
| LAST author removed | 16 |
| Abstract had only one author, additional author(s) added | 9 |
| Abstract had multiple authors, only FIRST author remained | 1 |
| FIRST author changed, LAST author remained unchanged | 33 (14.4) |
| FIRST author moved to be a middle author | 25 |
| FIRST author removed | 8 |
| FIRST and LAST authors swapped positions | 6 (2.6) |
| No change in middle authors | 1 |
| With at least one change in middle authors | 5 |
| LAST author changed to FIRST author | 15 (6.5) |
| FIRST author moved to be a middle author | 8 |
| FIRST author removed | 7 |
| FIRST author changed to LAST author | 15 (6.5) |
| LAST author moved to a middle author | 7 |
| LAST author removed | 5 |
| Abstract had only one author | 3 |
| Both FIRST AND LAST author changed | 37 (16.1) |
| Both authors moved to be middle authors | 14 |
| Both authors removed | 18 |
| FIRST author moved to be a middle author, LAST author removed | 2 |
| LAST author moved to be a middle author, FIRST author removed | 3 |

Table 4-3: Associations between conflicts of interest of authors of abstracts of RCTs presented at ARVO conferences during years 2001-2004 and likelihood of full publication of those abstracts, overall model (model 1) and interaction model, stratified by statistical significance of results for the main outcome (model 2)

| Author conflicts of interest | Model 1** | Model 2*** | | | |
|---|---|---|---|---|---|
| | RR (95% CI) | RR (95% CI) among abstracts with main outcome NOT statistically significant | RR (95% CI) among abstracts with main outcome STATISTICALLY SIGNIFICANT | RR (95% CI) among abstracts with statistical significance of main outcome NOT REPORTED | P-value of F-test of interaction **** |
| Conflicts of interest of **FIRST** author (years 2002-2004 only) | N=400 | N=87 | N=79 | N=234 | |
| At least one conflict of interest | **1.31 (1.04-1.64)** | 1.13 (0.75-1.72) | 1.22 (0.82-1.81) | 0.93 (0.63-1.37) | 0.7297 |
| F - Financial support* | **1.50 (1.19-1.90)** | 1.44 (0.95-2.17) | 1.37 (0.92-2.06) | 1.11 (0.74-1.67) | 0.7460 |
| I - Personal finance interest* | 1.17 (0.44-3.14) | - - | - - | - - | - |
| E - Employee of business with interest* | 0.54 (0.23-1.28) | 0.64 (0.20-2.03) | 0.96 (0.04-3.92) | 0.21 (0.03-1.38) | 0.6428 |
| C - Consultant of business with interest* | 1.10 (0.78-1.55) | 0.87 (0.41-1.86) | 1.08 (0.58-2.02) | 0.86 (0.52-1.40) | 0.6924 |
| P - Inventor/developer with patent* | 0.88 (0.36-2.16) | - - | - - | - - | - |
| R - Received gifts within the past year* | **1.42 (1.05-1.92)** | 1.47 (0.87-2.46) | 1.49 (0.97-2.28) | 0.91 (0.52-1.60) | 0.9221 |
| Conflicts of interest of **LAST** author (years 2002-2004 only) | N=400 | N=87 | N=79 | N=234 | |
| At least one conflict of interest | 1.06 (0.82-1.37) | 0.93 (0.59-1.47) | 1.06 (0.67-1.70) | 0.75 (0.50-1.14) | 0.8468 |
| F - Financial support* | 0.84 (0.53-1.34) | 0.73 (0.36-1.50) | 0.81 (0.34-1.97) | 0.58 (0.25-1.36) | 0.9637 |
| I - Personal finance interest* | - - | - - | - - | - - | - |
| E - Employee of business with interest* | 1.06 (0.77-1.47) | 0.89 (0.45-1.75) | 1.08 (0.58-2.02) | 0.82 (0.51-1.31) | 0.7582 |
| C - Consultant of business with interest* | 1.32 (0.86-2.03) | **1.73 (1.13-2.64)** | 1.56 (0.85-2.87) | 0.52 (0.15-1.76) | 0.3569 |
| P - Inventor/developer with patent* | 0.94 (0.32-2.75) | - - | - - | - - | - |
| R - Received gifts within the past year* | 1.18 (0.69-2.01) | 1.52 (0.83-2.78) | 1.01 (0.25-4.12) | 0.76 (0.30-1.91) | 0.7017 |
| Not reported | 1.18 (0.66-2.10) | - - | - - | - - | - |
| Not applicable (abstract had only one author) | 1.00 (0.54-1.86) | **1.64 (1.00-2.68)** | 1.28 (0.72-2.29) | 1.03 (0.50-2.12) | 0.6808 |

| Author conflicts of interest | Model 1** | Model 2*** | | | |
|---|---|---|---|---|---|
| | RR (95% CI) | RR (95% CI) among abstracts with main outcome NOT statistically significant | RR (95% CI) among abstracts with main outcome STATISTICALLY SIGNIFICANT | RR (95% CI) among abstracts with statistical significance of main outcome NOT REPORTED | P-value of F-test of interaction **** |
| Conflicts of interest of **ANY** author (years 2001-2004) | N=513 | N=111 | N=117 | N=285 | |
| At least one conflict of interest | 1.13 (0.93-1.38) | 1.01 (0.71-1.45) | 1.09 (0.77-1.56) | 0.81 (0.59-1.11) | 0.7158 |
| F - Financial support* | 1.16 (0.93-1.45) | 1.02 (0.68-1.53) | 1.18 (0.82-1.71) | 0.81 (0.55-1.18) | 0.8314 |
| I - Personal finance interest* | 0.74 (0.24-2.31) | - - | - - | - - | - |
| E - Employee of business with interest* | 1.07 (0.82-1.38) | 0.99 (0.62-1.58) | 1.13 (0.66-1.94) | 0.80 (0.55-1.17) | 0.8647 |
| C - Consultant of business with interest* | 1.18 (0.91-1.52) | 0.94 (0.52-1.69) | 1.25 (0.86-1.81) | 0.83 (0.55-1.25) | 0.6842 |
| P - Inventor/developer with patent* | 0.86 (0.43-1.71) | - - | - - | - - | - |
| R - Received gifts within the past year* | **1.35 (1.05-1.72)** | **1.53 (1.03-2.27)** | 1.36 (0.90-2.05) | 0.95 (0.64-1.42) | 0.8499 |

* More than one option could apply to each abstract.

** Model 1 – Overall model.

*** Model 2 – Interaction model – Results are stratified by whether the main outcome results were not statistically, statistically significant, or not reported.

**** F-test of interaction tests the overall statistical significance of the interaction between conflict of interest and all interaction terms in model 2. We added two interaction terms in model 2: (1) whether or not results for the main outcome in the abstract were statistically significant, and (2) whether or not statistical significance of results for the main outcome was reported in the abstract.

## FIGURES

Figure 4-1: Abstracts presented at ARVO conferences during years 2001-2004

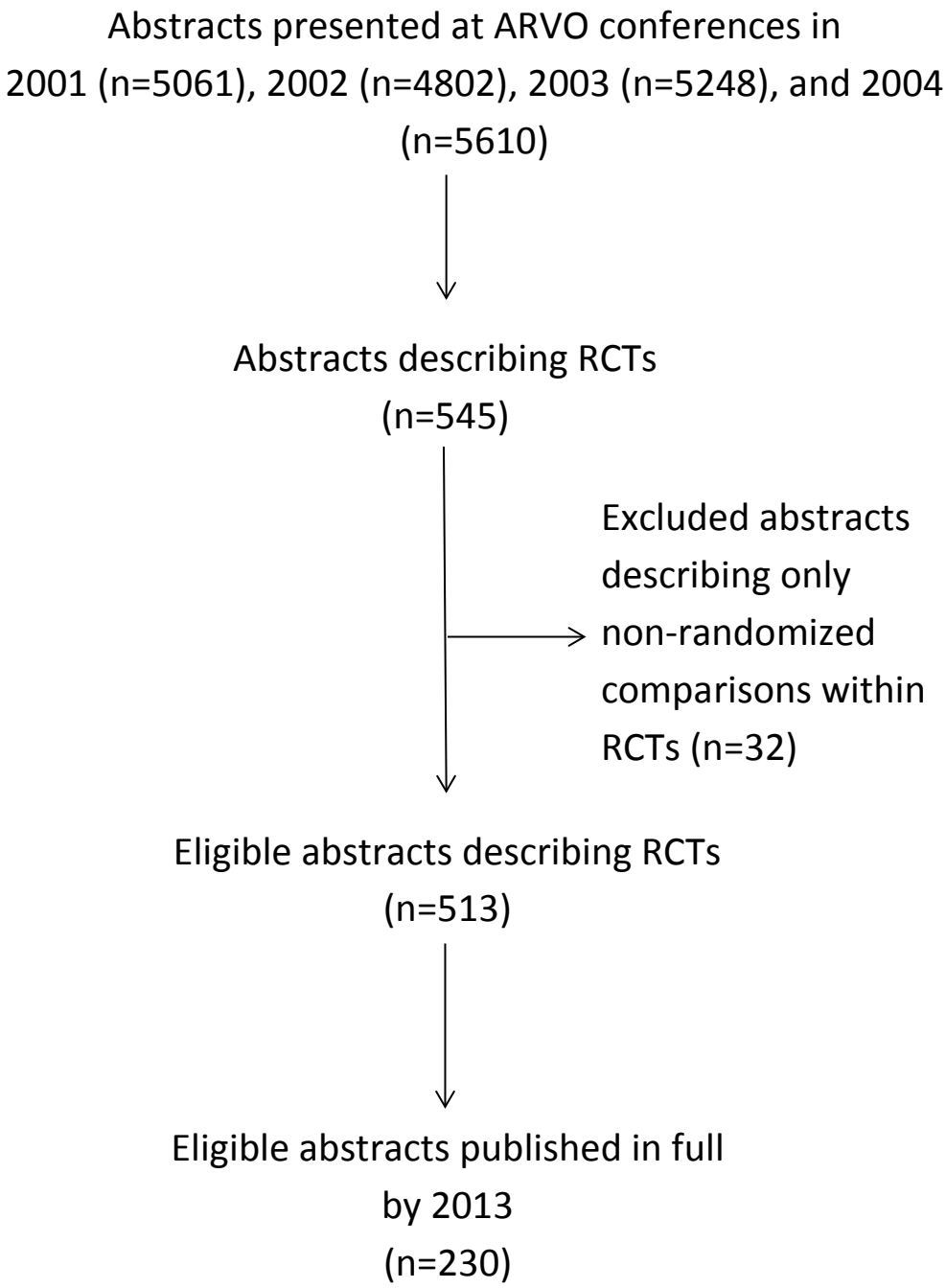


Figure 4-2: Kaplan-Meier plots showing time to full publication of abstracts of RCTs presented at ARVO conferences during years 2001-2004, overall and by various study characteristics, author characteristics, and statistical significance of main outcome results

Figure 4-2a: All abstracts

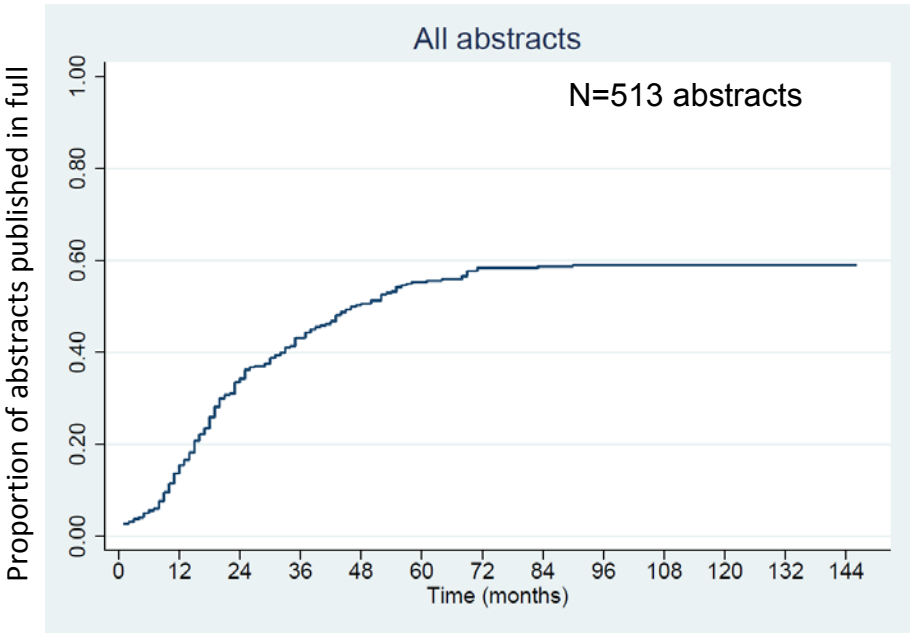
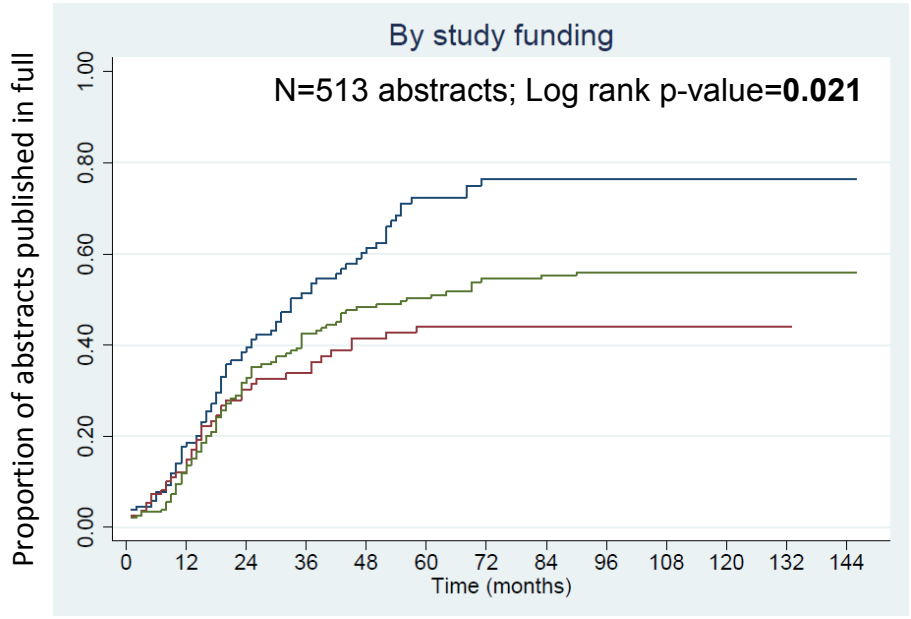
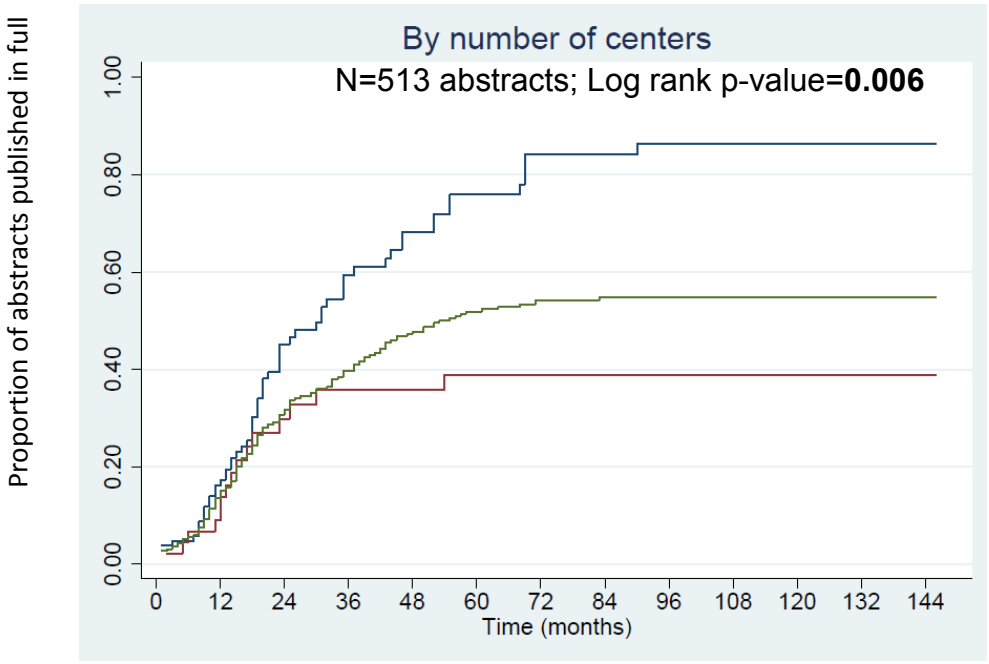



Figure 4-2b: By whether the RCT described in the abstract was funded, not funded, or funding was not reported



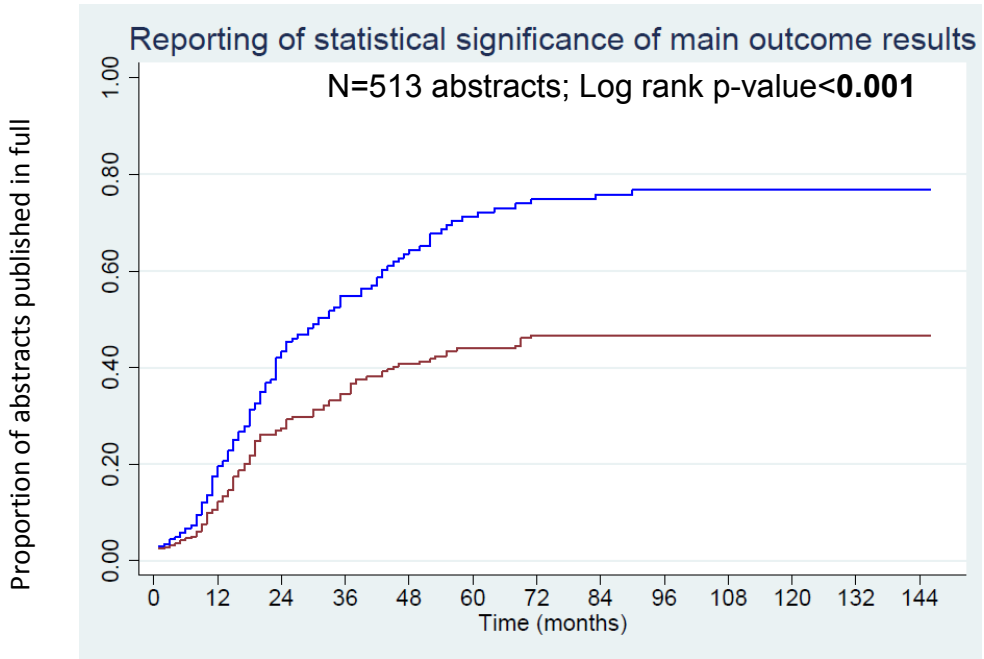

Legend: blue=funded; green=not reported; maroon=not funded

Figure 4-2c: By number of centers in the RCT described in the abstract



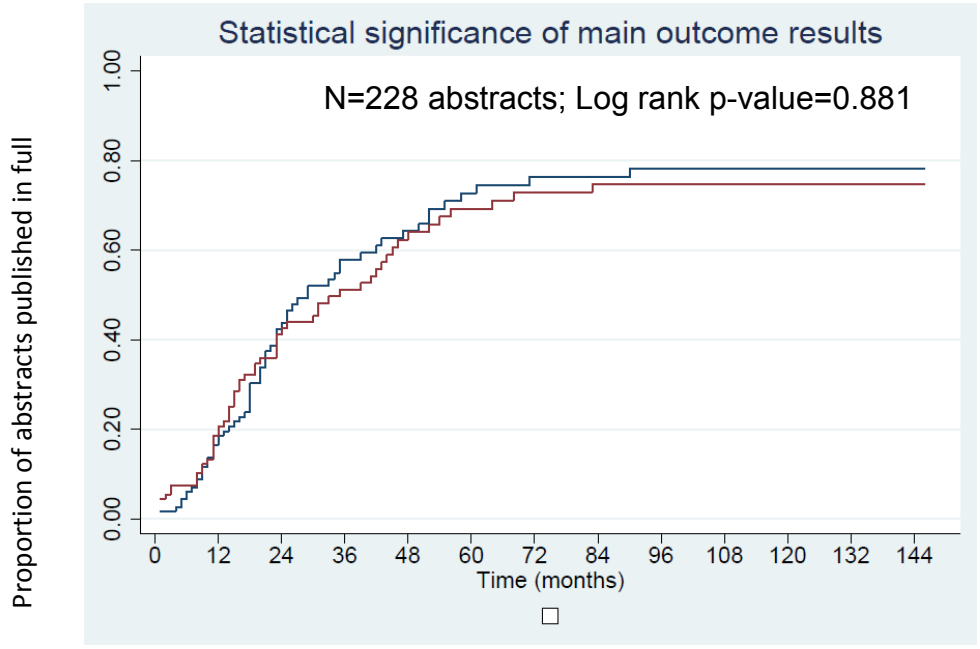

Legend: blue=multicenter; green=not reported; maroon=single center

Figure 4-2d: By whether or not statistical significance of results for the main outcome was reported in the abstract



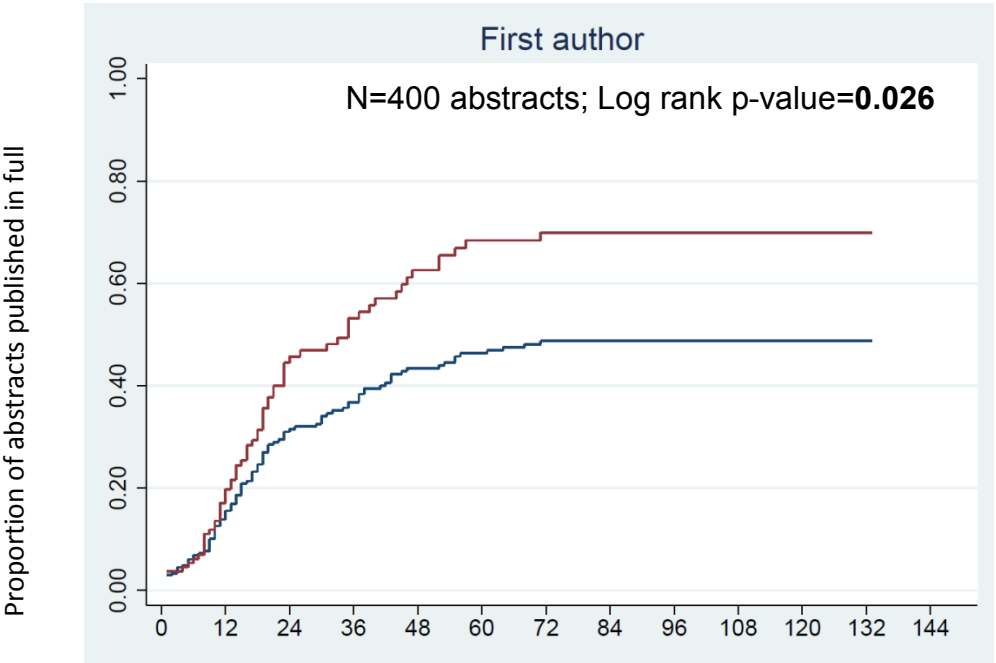

Legend: blue=reported; maroon=not reported

Figure 4-2e: By statistical significance of results for the main outcome for abstracts in which this was reported



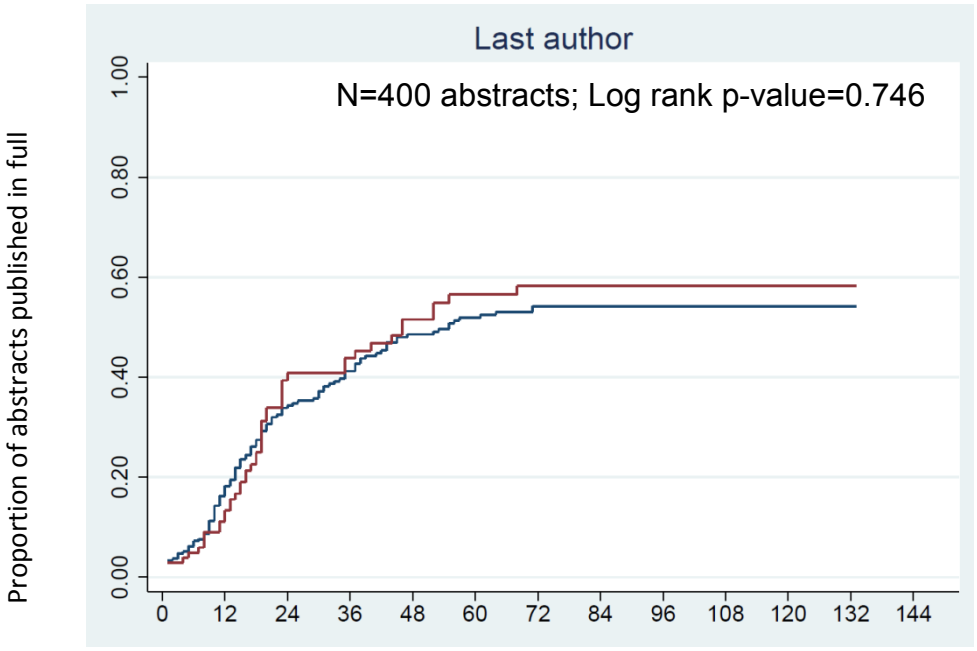

Legend: maroon=statistically significant; blue=not statistically significant

Figure 4-2f: By whether or not the FIRST AUTHOR of the abstract reported at least one conflict of interest (abstracts from 2002-2004 only)



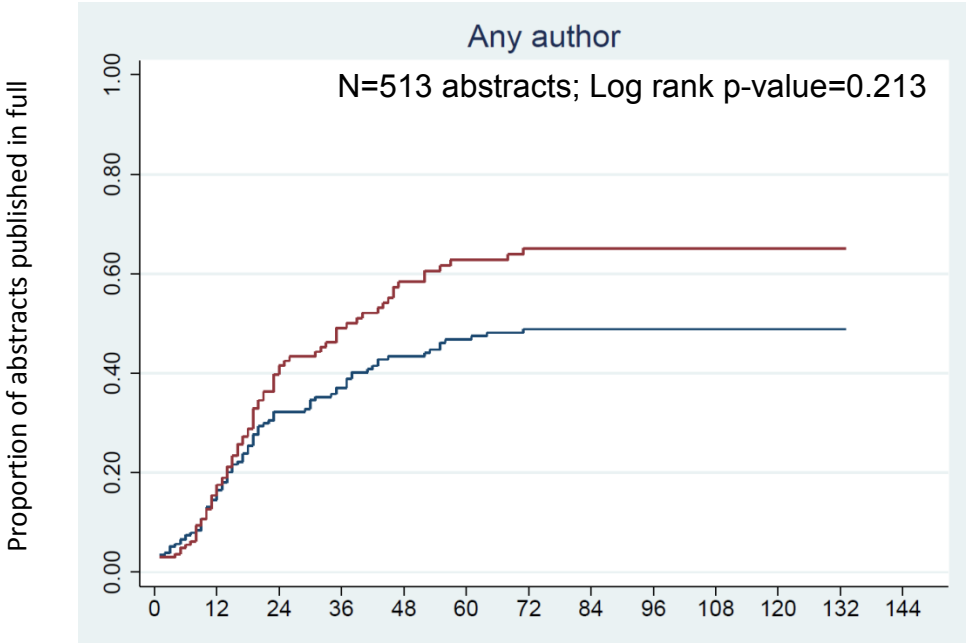

Legend: maroon=at least one conflict of interest; blue=no conflict of interest

Figure 4-2g: By whether or not the LAST AUTHOR of the abstract reported at least one conflict of interest (abstracts from 2002-2004 only)




Legend: maroon=at least one conflict of interest; blue=no conflict of interest

Figure 4-2h: By whether or not the ANY AUTHOR of the abstract reported at least one conflict of interest (abstracts from 2002-2004 only)




Legend: maroon=at least one conflict of interest; blue=no conflict of interest

Figure 4-3: Amount of agreement in main outcome results in 86 pairs of conference abstracts and full publications. Exact agreement (green), qualitative discordance (yellow), and various categories of quantitative discordance (blue) are depicted under two different definitions of agreement – exact agreement (left) and exact agreement or <10% discordance (right).

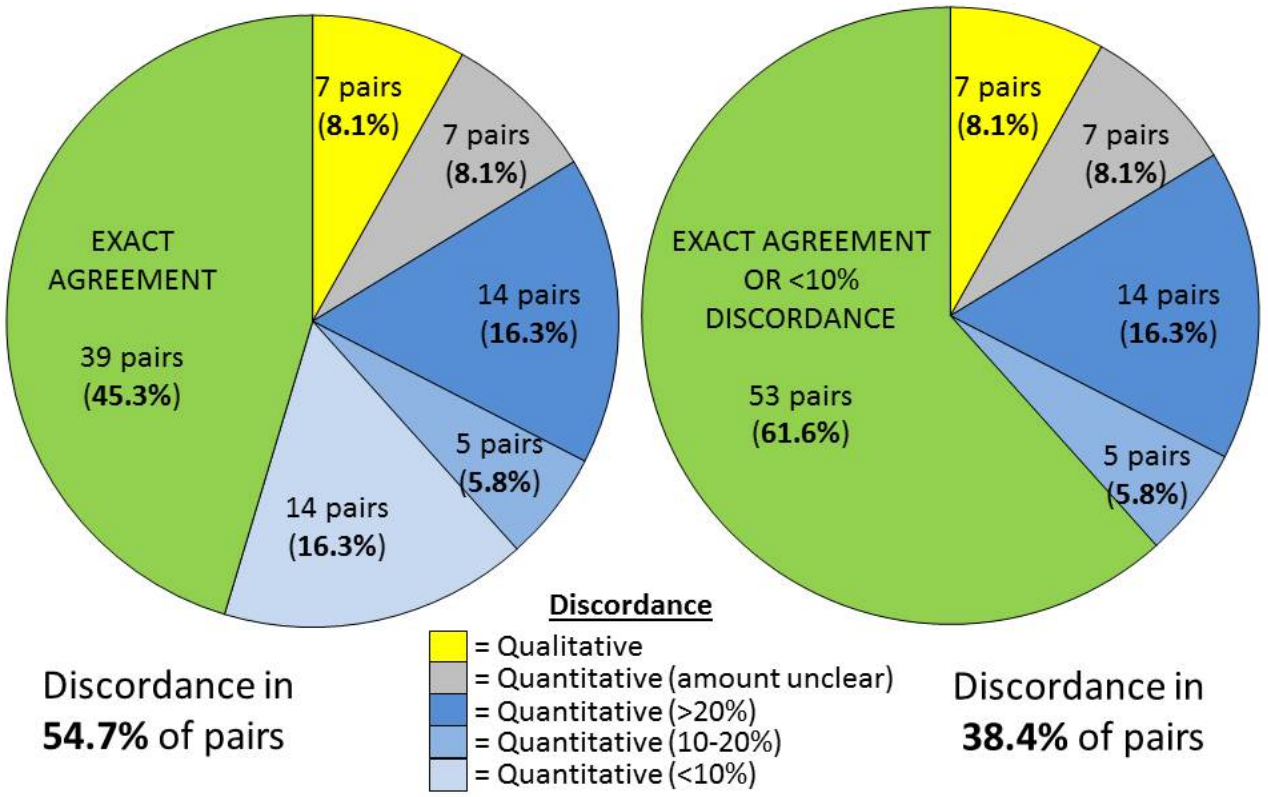
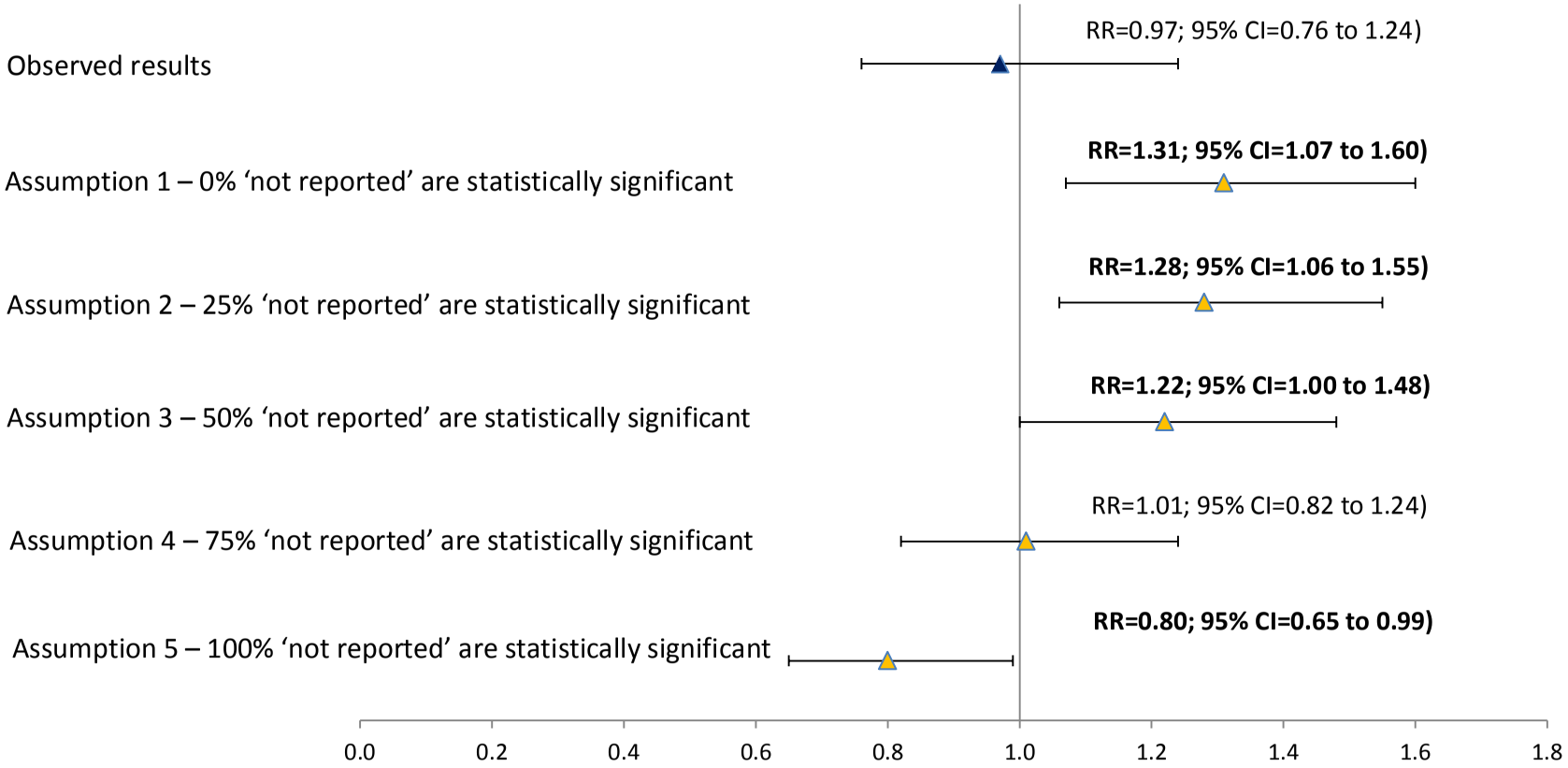


Figure 4-4: Unadjusted associations between statistical significance of results for the main outcome in abstracts of RCTs presented at ARVO conferences during years 2001-2004 and likelihood of full publication of those abstracts. Results of observed results and hypothetical analyses are presented under five different assumptions about the proportion of abstracts with statistical significance of results for the main outcome among abstracts not reporting statistical significance of results for the main outcome.



# CHAPTER FIVE

## Conclusions

***Summary***

A key premise of evidence-based healthcare is that, in addition to clinical expertise and patient values, healthcare should be guided by sound clinical intervention research.[1] In turn, sound clinical intervention research relies on the use of the right outcomes dependably and without bias. We viewed the process of outcome use in a clinical intervention research study (clinical trial or systematic review) as having four main steps: (1) selection, (2) specification, (3) data collection and analysis, and (4) outcome reporting. We designed this dissertation with three specific aims addressing important questions related to these steps. In this chapter, we summarize results and implications for each specific aim, make overall conclusions, and suggest future directions.

*Aim 1: Use social network analysis methods to (a) understand patterns of co-occurrence of outcomes in systematic reviews of HIV/AIDS; and (b) identify outcomes that are central to the network of outcomes examined in systematic reviews of HIV/AIDS.*

Drawing from the toolbox of social network analysis, we tested the novel application of those methods to systematically identify *central* outcomes from among the 294 unique outcomes in all 140 Cochrane systematic reviews in the field of HIV/AIDS. These central outcomes could be used to develop core outcome sets, which are valuable for the process of selecting outcomes for clinical intervention research.

Systematic reviews are an excellent starting point for identifying central outcomes. Existing systematic reviews in a field, when considered together, appraise a large portion of the evidence, much of it from clinical trials. Social network analysis capitalizes on the underlying affinity (or repulsion) between outcomes that leads to outcomes co-occurring (or not) in the same systematic review or clinical trial. Generally, the seven most central outcomes were not exactly the same as the seven most frequent outcomes, both overall as well as within topic subgroups of reviews defined by Cochrane. Using the field of HIV/AIDS as an example, we have shown that if frequency were used as the sole determinant for defining a core outcome set, certain important outcomes (such as adverse events) would be missed. Central outcomes and frequent outcomes should be used in tandem when selecting outcomes for a core outcome set.

*Aim 2: Assess the completeness of pre-specification and comparability of outcomes in systematic reviews addressing four common eye conditions.*

In clinical intervention research, *pre-specification* of outcomes occurs at the time the protocol is written, and *specification* of outcomes occurs when the results are reported. We examined all 57 Cochrane systematic review protocols addressing four common eye conditions and found that, if outcome pre-specification is judged using recommended standards for clinical trials, pre-specification of outcomes in these systematic review protocols is largely incomplete. We used a five-element framework (*domain*, *specific measurement*,

*specific metric*, *method of aggregation*, and *time-points*) to examine each outcome for completeness.

It is important to note that pre-specification is not an end in itself. Rather, it is important for two reasons. First, pre-specification helps ensure that the outcomes are measured and analyzed in an unbiased manner during the clinical trial. In our sample of Cochrane systematic review protocols, we are unable to rule out that incomplete pre-specification might have led the systematic reviewers to measure and/or analyze the outcomes differently or at different time-points than originally planned. If, indeed, these decisions were based on the systematic review results, it raises concerns about the validity of the analyses. Second, pre-specification of outcomes helps minimize the potential for outcome reporting bias. When the elements of outcomes pre-specified (i.e., before the study) and specified (i.e., when the study is reported) differ, this raises concerns about outcome reporting bias. The largely incomplete outcome pre-specification meant that we were unable to rule out that there might have been outcome reporting bias in these systematic reviews.

We encourage systematic reviewers and clinical trialists to incorporate all five elements when pre-specifying outcomes for their research, so that the subsequent steps of outcome use (data collection and analysis and outcome reporting) are conducted without bias.

Because of the incomplete pre-specification of outcomes, we were also unable to assess conclusively the comparability in individual outcome elements

across the various protocols per condition. However, we observed variation in specific metrics and methods of aggregation.

*Aim 3: Evaluate, using randomized controlled trials in ophthalmology, (a) the agreement in reported main outcome results comparing abstracts and their corresponding full publications; and (b) the association between author conflicts of interest and full publication of results presented in abstracts.*

We examined all 513 abstracts describing randomized controlled trials that were presented at the 2001-2004 Association for Research in Vision and Ophthalmology (ARVO) annual conferences, and followed them longitudinally for whether or not they were published in full through 2013. When we compared abstracts that were published in full with its corresponding full publications, we found that there was exact agreement in reported results for the main outcome in less than half of the abstract/full publication pairs. There was some amount of quantitative discordance in almost half the pairs, and qualitative discordance in 1/12 pairs. We also found that, irrespective of main outcome results, the conflicts of interest most associated with full publication of abstracts of RCTs are conflicts of interest of the abstract's first author (specifically, receiving financial support and at least one gift from industry in the past year).

***Overall conclusions and future directions***

The results we present in this dissertation provide numerous causes for concern, and we propose suggestions for ways forward.

The clinical research community is interested in finding the best method to develop core outcome sets for intervention research. A variety of semi-structured group discussions (e.g., workshops), unstructured group discussions, literature reviews, and surveys have been used, but no best practice method to do so currently exists.[2] Our results using social network analysis explored a new application to assessment of possible outcomes for a core outcome set (i.e., co-occurrence of outcomes in systematic reviews), leading to the identification of central outcomes for clinical intervention research. We propose that this method be examined in other fields and, if found to be valuable in developing core outcome sets, subsequently implemented more widely.

Specific questions remain. First, it is unclear to what extent the most central outcomes in systematic reviews (as identified in Aim 1) reflect the most central outcomes in clinical trials. It is likely that the number of outcomes examined in the clinical trials that were included in the systematic reviews we examined is greater than the number of outcomes examined in the systematic reviews alone. However, we currently do not know about the patterns of co-occurrence of outcomes among the clinical trials and between the clinical trials and the systematic reviews. A comparison of the social networks of outcomes in clinical trials and systematic reviews, and the overlap between the two lists of central outcomes obtained

therefrom, would contribute to our understanding of whether and to what extent core outcome sets would differ based on the source used for identifying central outcomes. It is possible, even likely, that the lists of the most central outcomes in systematic reviews and clinical trials would differ, given the different purposes of these two study designs (clinical trials are more focused on clinically-relevant outcomes, while systematic reviews are more focused on policy-relevant outcomes). Second, we found that the lists of the most central outcomes and the most frequent outcomes tended to differ more in networks of outcomes that were more centralized (higher *centralization* values, see Chapter 2). This suggests that the social network analysis approach might be more valuable in more centralized networks compared to less centralized networks (i.e., networks in which the overlap between central and frequent outcomes is greater). Future studies in other fields should explore this finding further. If it is indeed to found to be generally true that the social network analysis method might be more valuable in more centralized networks, the degree of centralization could be used to determine whether a complete social network analysis of that network is worth the resources spent in conducting and interpreting the analysis.

Our results related to outcome specification also provide cause for concern. Incomplete pre-specification of outcomes suggests that we are unable to rule out the possibility of bias in the data collection and analysis in these systematic reviews. Incomplete specification also hinders comparability in outcomes across systematic reviews (or clinical trials) addressing similar interventions in a field.

That systematic reviewers do not completely pre-specify outcomes they will examine is worrisome because of the potential for outcome reporting bias. This could occur if the systematic reviewer's choice of outcomes is influenced by the outcomes examined (and the results) in the clinical trials that would be included in the systematic review. Future research should evaluate whether the results of outcomes examined in the clinical trials influence choice of outcomes in systematic reviews.

We encourage (1) organizations that provide methodological and reporting guidance for researchers conducting systematic reviews and clinical trials (e.g., Cochrane, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), and Consolidated Standards of Reporting Trials (CONSORT)), to incorporate the five-element framework for outcome specification into guidance documents; and (2) journal editors and peer reviewers to insist on appropriate reporting of all outcomes using this framework.

Our results related to discrepancies in main outcome results between conference abstracts and their corresponding full publications are a cause for concern. Our results indicate that more than half of abstract/full publication pairs have some form of discordance for the main outcome. We encourage all clinical trialists to be vigilant about reporting dependable clinical trial data, whatever the forum of presentation. While we support the inclusion of results from abstracts in systematic review, we encourage systematic reviewers to be aware of the fact that

discordances can occur, and to conduct the appropriate sensitivity analyses to examine the robustness of their findings.

Another concern that our results raise is that, irrespective of statistical significance of study results, monetary gain (conflicts of interest) might determine whether authors of a conference abstract describing a randomized controlled trial might publish the abstract in full. The first author's conflict of interest appears to matter most, specifically, receiving financial support or gifts from industry within the past year. We explored certain potential reasons why these findings might have been observed. First, compared with last authors, first authors were less often employed by industry and more often affiliated with academic institutions, where publications generally help towards career advancement. Second, compared with last authors, first authors more often received financial support and gifts from industry, which might have specifically encouraged first authors to devote the necessary time and resources to publish. Third, it is possible that the authors in other study were different from authors in previous studies; for example, the proportion of authors reporting at least one conflict of interest was lower than what has been documented in previous studies.[3, 4]

Finally, less than half of the abstracts we included were published in full. Because of the importance of publication and the commitment made by clinical trialists to clinical trial participants to advance science, we encourage all clinical trialists to publish their results.

## REFERENCES


1. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based medicine: How to practice and teach EBM*. 2nd edition. 2000.

2. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, Williamson PR. Choosing health outcomes for comparative effectiveness research: A systematic review. *PLOS One* 2014. DOI: 10.1371/journal.pone.0099111.

3. Campbell EG, Louis KS, Blumenthal D. Looking a gift horse in the mouth: corporate gifts support life sciences research. *JAMA* 1998; 279:995-9.

4. Neuman J, Korenstein D, Ross JS, Keyhani S. Prevalence of financial conflicts of interest among panel members producing clinical practice guidelines in Canada and United States: cross sectional study. BMJ 2011. Oct 11;343:d5621. doi: 10.1136/bmj.d5621.

# CURRICULUM VITAE

**Ian Jude Saldanha, MBBS, MPH, PhD**
Curriculum Vitae, July 23, 2015

# PART I

## CONTACT INFORMATION

Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe Street, Room E6006
Baltimore, MD 21205, USA
E-mail: isaldan1@jhmi.edu
Phone: +1-443-603-4684

## EDUCATION AND TRAINING

### Degrees

| Year | Degree | Institution | Field |
|---|---|---|---|
| 2015 | Doctor of Philosophy (PhD) | Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA | Epidemiology (Clinical Trials and Evidence Synthesis) |
| 2007 | Master of Public Health (MPH) | Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA | Public Health (Child and Adolescent Health) |
| 2006 | Bachelor of Medicine and Bachelor of Surgery (MBBS) | Grant Medical College, Mumbai, India | Medicine and Surgery |

### PhD Dissertation

Title: Outcomes in clinical intervention research: Selection, specification, and reporting

Advisor: Dr. Kay Dickersin

**PROFESSIONAL EXPERIENCE**

09/2011-Current *Graduate Research Assistant*
Cochrane Eyes and Vision Group US Project, Baltimore, Maryland, USA
Lead, implement, and manage a methodological project examining factors associated with full publication of results of randomized trials, and data discrepancies between conferences abstracts and publications.

02/2010–09/2011 *Project Manager*
Johns Hopkins Evidence-based Practice Center (EPC), Baltimore, Maryland, USA
Designed, implemented, and managed methodological projects and evidence syntheses; for the Agency for Healthcare Research and Quality (AHRQ). Projects focused on methods for systematic reviews and meta-analyses and prioritizing research needs.

10/2007–09/2011 *Senior Research Program Coordinator*
General Internal Medicine, Johns Hopkins University, Baltimore, Maryland, USA
Designed, implemented, and managed systematic reviews and meta-analyses in collaboration with the Cystic Fibrosis Foundation to inform their evidence-based clinical practice guidelines; wrote manuscripts for publication in peer-reviewed journals; managed team members and deadline-oriented projects; and developed and maintained relational databases.

02/2007–10/2007 G*raduate Research Assistant*
General Internal Medicine, Johns Hopkins University, Baltimore, Maryland, USA
Collected and analyzed data for systematic reviews and meta-analyses in cystic fibrosis.

09/2006–01/2007 *Graduate Research Assistant*
Health Sciences Informatics, Johns Hopkins University, Baltimore, Maryland, USA
Collected data for a systematic review and meta-analysis in developmental dysplasia of the hip (DDH).

02/2006–05/2006 *Assistant Physician and Research Assistant*
Human Healthcare and Research Foundation, Mumbai, India
Recruited patients, conducted clinical assessments, performed data entry, and managed a database for a clinical trial evaluating interventions to improve adherence to highly active antiretroviral therapy (HAART) in HIV/AIDS.

01/2005–12/ 2005 *Medical Intern*
Grant Medical College and Sir JJ Group of Hospitals, Mumbai, India
Preventive and social medicine, internal medicine, general surgery, obstetrics and gynecology, pediatrics, orthopedics, emergency medicine, pathology, dermatology and venereal diseases, anesthesiology, ophthalmology, and otorhinolaryngology.

## PROFESSIONAL ACTIVITIES

### Society Membership

2005-Current Maharashtra Medical Council, India

2011-2012 Society for Clinical Trials (SCT)

2005 Flood Relief Disaster Management Team, Government of Maharashtra, Mumbai, India

### Society Leadership

2014-2015 *Epidemiology Student Organization Student Representative to Department of Epidemiology Faculty Meetings*
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2013-2014 *President*, Epidemiology Student Organization
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2012-2013 *Teaching Assistant Training Chair*, Epidemiology Student Organization
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2005-2006 *Executive Committee Member,* Catholic Medical Guild of St. Luke, Mumbai, India

2005 *Monitor,* World Health Organization National Polio Surveillance Program, Mumbai, India. Managed and supervised over 50 teams of multi-disciplinary vaccinators to help ensure comprehensive and effective coverage of polio vaccination.

**Meetings Organized**

2013-2014 & 2011-2012 Clinical Trials and Evidence Synthesis Journal Club
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2013 Missing Data Mini-Symposium (Visiting scholar - Dr. Jim Neaton)
Center for Clinical Trials and Evidence Synthesis, Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2013 Department of Epidemiology Teaching Assistant Training Session
Epidemiology Student Organization
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2012 Visiting Scholar Days (Visiting scholar - Dr. Robert Califf)
Center for Clinical Trials and Evidence Synthesis, Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

**Other**

2012-2015 *Student Mentor*
Sheriza Baksh, Gillian Gresham, Jimmy Le, and Eric Rubenstein
Epidemiology Student Organization
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

## EDITORIAL ACTIVITIES

### Peer Review Activities

1. *American Journal of Epidemiology*
2. *Cochrane Database of Systematic Reviews* - Cochrane Cystic Fibrosis and Genetic Disorders Group
3. *Agency for Healthcare Research and Quality (AHRQ)* - Effective Health Care (EHC) Program
4. *Journal of Clinical Epidemiology*
5. *Journal of Cystic Fibrosis*
6. *Ophthalmology*
7. *Pediatric Pulmonology*
8. *Pediatrics*
9. *PLoS Medicine*
10. *PLoS One*
11. *Research Synthesis Methods*
12. *Systematic Reviews*
13. *Trials*

### Abstracts Committees

1. *22nd Cochrane Colloquium, Hyderabad, India 2014*

## HONORS AND AWARDS

### Honors

2014 *Thomas C. Chalmers Award for Best Oral Presentation*
22nd Cochrane Colloquium, Hyderabad, India

2014 *Miriam E. Brailey Fund Award in Epidemiology*
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2014 & 2013 *Teaching Assistant Recognition Award for Outstanding Service* (two times)
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2013 *Dorothy and Arthur Samet Student Support Fund Award in Epidemiology*
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

**Awards**

2015 *Student Travel Support Fund Award*
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2012-2015 *Student Travel Award* (four times)
Center for Clinical Trials and Evidence Synthesis
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2014 *Open Access Promotion Fund Award*
Johns Hopkins Libraries, Baltimore, USA

2014 *Center for Global Health Conference Grant Award*
Johns Hopkins Center for Global Health, Baltimore, Maryland, USA

2013 *Student Conference Fund Award*
Student Assembly
Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

2013 *Global Health Established Field Placement Site Award*
Collaborative Outcomes Study of Meta-Analyses in HIV/AIDS, Cape Town, South Africa
Johns Hopkins Center for Global Health, Baltimore, Maryland, USA

## PEER-REVIEWED PUBLICATIONS

### Journal Articles

1. **Saldanha IJ**, Li T, Yang C, Ugarte-Gil C, Rutherford GW, Dickersin K. Use of social network analysis to identify central outcomes for clinical research: A case study of systematic reviews of HIV/AIDS. (Submitted)

2. **Saldanha IJ**, Dickersin K, Wang X, Li T. Outcomes in Cochrane systematic reviews addressing four common eye conditions: An evaluation of completeness and comparability. *PLoS One* 2014; 9(10): e109400.

3. Li T, **Saldanha IJ**, Vedula SS, Yu T, Rosman L, Twose C, Goodman S, Dickersin K. Learning by doing – Teaching systematic review methods in 8 weeks. *Res Synth Meth* 2014. DOI: 10.1002/jrsm.1111.

4. **Saldanha IJ**, Akinyede O, McKoy NA, Robinson KA. Immunosuppressive drug therapy for preventing rejection following lung transplantation in cystic fibrosis. *Cochrane Database Syst Rev* 2013, Issue 12. Art. No.: CD009421. DOI: 10.1002/14651858.CD009421.pub2.
   - **Saldanha IJ**, McKoy NA, Robinson KA. Immunosuppressive drug therapy for preventing rejection following lung transplantation in cystic fibrosis (Protocol). *Cochrane Database Syst Rev* 2011, Issue 11. Art. No.: CD009421. DOI: 10.1002/14651858.CD009421.

5. Robinson KA, Odelola OA, **Saldanha IJ**, McKoy N. Palivizumab for prophylaxis against respiratory syncytial virus infection in children with cystic fibrosis. *Cochrane Database Syst Rev* 2013; 6:CD007743. doi: 10.1002/14651858.CD007743.pub4.
   - Robinson KA, Odelola OA, **Saldanha I**, McKoy N. Palivizumab for prophylaxis against respiratory syncytial virus infection in children with cystic fibrosis (Protocol). *Cochrane Database Syst Rev 2009*, Issue 2. Art. No.: CD007743. DOI: 10.1002/14651858.CD007743.

6. **Saldanha IJ**, Wilson LM, Bennett WL, Nicholson WK, Robinson KA. Development and pilot test of a process to identify research needs from a systematic review. *J Clin Epidemiol* 2013; 66(5) 538-545.

7. Robinson KA, McKoy N, **Saldanha IJ**, Odelola OA. Active cycle of breathing technique for cystic fibrosis. *Cochrane Database Syst Rev* 2012; 12:CD007862. doi: 10.1002/14651858.CD007862.pub3.
   - Robinson KA, McKoy N, **Saldanha I**, Odelola OA. Active cycle of breathing technique for cystic fibrosis (Protocol). *Cochrane Database Syst Rev* 2009, Issue 3. Art. No.: CD007862. DOI: 10.1002/14651858.CD007862.

8. Bennett WL, Robinson KA, **Saldanha IJ**, Wilson LM, Nicholson WK. High priority research needs for gestational diabetes mellitus. *J Womens Health (Larchmt)* 2012 Sep; 21(9):925-32.

9. Robinson KA, **Saldanha IJ**, McKoy NA. Development of a framework to identify research gaps from systematic reviews. *J Clin Epidemiol* 2011; 64(12):1325-30.

10. Li H, Robinson KA, Anton B, **Saldanha IJ**, Ladenson PW. Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab* 2011; 96(11):E1719-26.

11. Robinson KA, **Saldanha IJ**, McKoy NA. Identification of research gaps from evidence-based guidelines: A pilot study in cystic fibrosis. *Int J Technol Assess Health Care* 2011; 27(3):247-252.

12. Robinson KA, **Saldanha IJ**, McKoy NA. Management of infants with cystic fibrosis: A summary of the evidence for the Cystic Fibrosis Foundation Working Group on Care of Infants with Cystic Fibrosis. *J Pediatr* 2009; 155(6 Suppl):S94-S105.

13. Golden SH, Robinson KA, **Saldanha I**, Anton B, Ladenson PW. Clinical review: Prevalence and incidence of endocrine and metabolic disorders in the United States: A comprehensive review. *J Clin Endocrinol Metab* 2009; 94(6):1853-1878.

**Peer-Reviewed Reports**

14. Robinson KA, Saldanha IJ, McKoy NA. Frameworks for determining research gaps during systematic reviews. *Agency for Healthcare Research and Quality* (Rockville, Maryland, USA); 2011 Jun. Report No.: 11-EHC043-EF.

15. Bennett WL, Nicholson WK, **Saldanha IJ**, Wilson LM, McKoy NA, Robinson KA. Future research needs for the management of gestational diabetes. *Agency for Healthcare Research and Quality* (Rockville, Maryland, USA); 2010 Nov. Report No.: 11-EHC005-EF.

## PART II

## TEACHING

### Lead Teaching Assistant

2012-2015 *Systematic Reviews and Meta-Analysis* (Lead Teaching Assistant) (four times)
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Drs. Kay Dickersin and Tianjing Li

### Teaching Assistant

2015 *Introduction to Systematic Reviews and Meta-Analysis*
Coursera
Faculty – Drs. Tianjing Li and Kay Dickersin

2013-2015 *Introduction to Systematic Reviews and Meta-Analysis* (three times)
Graduate Summer Institute of Epidemiology and Biostatistics
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Drs. Tianjing Li and Kay Dickersin

2013-2015 *Critical Reading of the Epidemiologic Literature* (three times)
Graduate Summer Institute of Epidemiology and Biostatistics
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Dr. Moyses Szklo

2014 *Methods for Clinical and Translational Research*
Graduate Summer Institute of Epidemiology and Biostatistics
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Dr. Jonathan Samet

2012 *Epidemiology Research Methods (340.752)*
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Drs. Gypsyamber D'Souza, Milo Puhan, and Stephan Ehrhardt

2012 *Introduction to Clinical Trials*
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Drs. Janet Holbrook and Lea Drye

2012 *Principles of Epidemiology*
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Drs. Rosa Crum and Gregory Kirk

**Guest Lecturer**

2013-2015 *Design and Conduct of Clinical Trials* (three times)
Graduate Summer Institute of Epidemiology and Biostatistics
Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
Faculty – Mr. David Shade and Dr. Elizabeth Sugar

2013 *Systematic Reviews and Meta-Analyses*
Stellenbosch University, Cape Town, South Africa
Faculty – Drs. Taryn Young and Tamara Kredo

**Workshop Instruction**

2015 & 2014 *Ensuring that Patient-Centered Outcomes are Included in Research*
Cochrane Colloquia, Vienna, Austria (2015) & Hyderabad, India (2014)

2012-2015 *Completing a Cochrane Systematic Review* (seven times)
Cochrane Eyes and Vision Group, Baltimore, Maryland, USA

2014 S*hould Cochrane Limit the Number of Outcomes in a Systematic Review?*
2014 Cochrane Colloquium, September 22-26, Hyderabad, India

## RESEARCH GRANT PARTICIPATION

### Grants, Current

October 1, 2014 – September 30, 2017
*Develop, Test and Disseminate a New Technology to Modernize Data Abstraction in Systematic Reviews*
Grant number: ME-1310-07009
Sponsoring Agency: Patient-Centered Outcomes Research Institute
PI: Dr. Tianjing Li
Role: *Project Director*
Objective: To develop a new software application, Data Abstraction Assistant (DAA), for data abstraction in systematic reviews, to evaluate DAA against two traditional data abstraction approaches, and to generate critical evidence needed for modernizing the systematic review process to improve efficiency and validity.

May 1, 2010 – April 30, 2017
*Comparative Effectiveness Research & Cochrane Eyes and Vision Group*
Grant number: 1 U01 EY020522
Sponsoring Agency: National Eye Institute, National Institutes of Health
PI: Dr. Kay Dickersin
Role: *Graduate Research Assistant*
Objective: To continue to serve as a coordinating center for comparative effectiveness research, specifically comparative effectiveness research related to systematic reviews, in eyes and vision in the US.

### Grants, Completed

February 1, 2007 – September 7, 2011
*Development of Evidence-based Guidelines*
Sponsoring Agency: Cystic Fibrosis Foundation
PI: Dr. Karen Robinson
Role: *Senior Research Program Coordinator*
Objective: To perform systematic reviews and provide consultation to facilitate guideline development.

October 1, 2007 – September 30, 2008
*Endocrine and Metabolism Epidemiology Database (EMED)*
Sponsoring Agency: The Endocrine Society, Journal of Clinical Endocrinology and Metabolism
PI: Dr. Sherita Golden
Role: *Senior Research Program Coordinator*
Objective: To develop EMED and related manuscripts.

**Contracts, Completed**

October 26, 2009 - October 25, 2012
*ARRA Comprehensive Comparative Effectiveness Reviews, Johns Hopkins Evidence-based Practice Center*
Sponsoring Agency: Agency for Healthcare Research and Quality (AHRQ)
PI: Dr. Eric Bass
Role: *Senior Research Program Coordinator*
The primary mission of the Evidence-based Practice Center for this award is to conduct topic refinements and comparative effectiveness reviews about important clinical topics.
1. Pilot Project on Identifying Research Needs: Identifying Research Needs in Gestational Diabetes Mellitus
2. Research Needs Methods Projects: Frameworks for determining research gaps (question #1)

**PRESENTATIONS**

**International**

1. **Saldanha IJ**, Li T, Yang C, Ugarte-Gil C, Rutherford GW, Dickersin K. Social network analysis for identifying central outcomes for clinical research: A case study using Cochrane reviews of HIV/AIDS (*Oral*)
   2015 Cochrane Colloquium, October 3-7, Vienna, Austria

2. **Saldanha IJ**, Scherer RW, Dickersin K. Agreement in results data between conference abstracts and full reports of randomized controlled trials: Should we depend on conference abstracts? (*Oral*)
   2015 Cochrane Colloquium, October 3-7, Vienna, Austria

3. Dickersin K, **Saldanha IJ**, Le J, Law A, Scherer RW, Li T. Use of a well-known surrogate outcome instead of a patient-important outcome can be viewed as research waste: Examination of an ad hoc sample of clinical trials and systematic reviews. (*Oral*)
   2015 REWARD / EQUATOR Conference, September 28-30, Edinburgh, Scotland

4. Scherer RW, **Saldanha IJ**, Parlett L, Dickersin K. Do trial registers close the gap in finding RCT results reported in conference abstracts? (*Oral*)
   2015 REWARD / EQUATOR Conference, September 28-30, Edinburgh, Scotland

5. **Saldanha IJ**, Li T, Yang C, Ugarte-Gil C, Rutherford GW, Dickersin K. Can we use social network analysis to identify central outcomes for clinical research? A case study using Cochrane reviews and clinical trials of HIV/AIDS. (*Oral*)
   2015 Society for Research Synthesis Methodology Annual Meeting, July 8-10, Nashville, Tennessee, USA

6. **Saldanha IJ**, Scherer RW, Rouse BD, Dickersin K. Impact of author conflicts of interest on the likelihood of full publication of randomized trials presented as conference abstracts. (*Poster*)
   2015 Society for Clinical Trials Annual Meeting, May 17-20, Arlington, Virginia, USA

7. **Saldanha IJ**, Li T, Heyward JS, Dickersin K. Are Cochrane review protocols available and post-protocol changes to methods documented? A study in HIV/AIDS and ophthalmology. (*Poster*)
   2014 Cochrane Colloquium, September 22-26, Hyderabad, India

8. **Saldanha IJ**, Dickersin K, Ugarte-Gil C, Li T, Rutherford G, Volmink J. Are we measuring enough of what patients want? A collaborative study of Cochrane reviews on HIV/AIDS. (*Oral*)
   2014 Cochrane Colloquium, September 22-26, Hyderabad, India

9. Fusco N, **Saldanha IJ**, Gresham G, Li T, Lack of originality in non-Cochrane reviews. (*Poster*)
   2014 Cochrane Colloquium, September 22-26, Hyderabad, India

10. **Saldanha IJ**, Ugarte-Gil C, Li T, Dickersin K, Rutherford G. Choosing the best outcomes when designing clinical trials: A case study using Cochrane reviews addressing HIV/AIDS. (*Oral*)
    2014 Society for Clinical Trials Annual Meeting, May 18-21, Philadelphia, Pennsylvania, USA

11. **Saldanha IJ**, Wang X, Li T, Dickersin K. Completeness of outcome specification across Cochrane systematic reviews of three common eye conditions: Time to be more explicit! (*Poster*)
    2013 Cochrane Colloquium, September 19-23, Quebec City, Canada

12. **Saldanha IJ**, Wang X, Li T, Dickersin K. Variation in outcome measure usage across Cochrane systematic reviews related to three common eye conditions. (*Oral*)
    2013 Cochrane Colloquium, September 19-23, Quebec City, Canada

13. Rosman L, Twose C, Li M, Li T, **Saldanha IJ**, Dickersin K. Teaching searching in an intensive systematic review course: "How many citations should I expect to review?" (*Poster*)
    2013 Cochrane Colloquium, September 19-23, Quebec, Canada

14. Twose C, Rosman, L, Gross P, Hesson D, Adamo J, Li T, **Saldanha I**, Vedula S, Dickersin K. An interdisciplinary collaboration to teach systematic review methods. (*Oral*)
    2013 Medical Library Association Annual Meeting, May 3-8, Boston, Massachusetts, USA

15. **Saldanha IJ**, Vedula SS, Yu T, Rosman L, Twose C, Li T, Dickersin K. Learning by doing - Teaching systematic review methods in 8 weeks. (*Poster*)
    2012 Cochrane Colloquium, September 30-October 3, Auckland, New Zealand

16. Agbedia OO, **Saldanha IJ**, Donath E, Chandra N, Robinson KA. Correlation between vitamin D and lung function in cystic fibrosis: Results of a systematic review. (*Poster*)
    25th Annual North American Cystic Fibrosis Conference, November 3-5, 2011, Anaheim, California, USA. *Pediatr Pulmonol* Suppl 2011; 34:400-1. Abstract Number 517.

17. **Saldanha IJ**, Agbedia OO, Donath E, Robinson KA. Supplementation with ergocalciferol (vitamin D2) versus cholecalciferol (vitamin D3) in cystic fibrosis: Results of a systematic review. (*Poster*)
    25th Annual North American Cystic Fibrosis Conference, November 3-5, 2011, Anaheim, California, USA. *Pediatr Pulmonol* Suppl 2011; 34:400. Abstract Number 516.

18. **Saldanha IJ**, Agbedia OO, Donath E, Robinson KA. Use of hydroxylated vitamin D supplements in cystic fibrosis: Results of a systematic review. (*Poster*)
    25th Annual North American Cystic Fibrosis Conference, November 3-5, 2011, Anaheim, California, USA. *Pediatr Pulmonol* Suppl 2011; 34:400. Abstract Number 515.

19. Palamaner Subash Shantha G, **Saldanha IJ**, Agbedia OO, Chandra N, Donath E, Robinson KA. Effect of ultraviolet therapy on vitamin D levels in patients with cystic fibrosis: Results of a systematic review. (*Poster*)
    25th Annual North American Cystic Fibrosis Conference, November 3-5, 2011, Anaheim, California, USA. *Pediatr Pulmonol* Suppl 2011; 34:399-400. Abstract Number 514.

20. **Saldanha IJ**, Robinson KA, McKoy NA**.** Palivizumab for prophylaxis against respiratory syncytial virus (RSV) infection in children with cystic fibrosis: results of a Cochrane systematic review. (*Poster*)
    24th Annual North American Cystic Fibrosis Conference, October 21-23, 2010, Baltimore, Maryland, USA. *Pediatr Pulmonol* Suppl 2010; 34:331. Abstract Number 313.

21. McKoy NA, **Saldanha IJ**, Robinson KA**.** A comparison of the active cycle of breathing technique (ACBT) to other airway clearance therapies in patients with cystic fibrosis. (*Oral*)
    24th Annual North American Cystic Fibrosis Conference, October 21-23, 2010, Baltimore, Maryland, USA. *Pediatr Pulmonol* Suppl 2010; 34:379-80. Abstract Number 446.

22. **Saldanha IJ**, Robinson KA, McKoy NA. Setting priorities in clinical research: identification and classification of research gaps from evidence-based guidelines. (*Oral*)
    2010 Guidelines International Network Conference, August 25-28, Chicago, Illinois, USA. *Otolaryngology - Head and Neck Surgery* 2010; 143 (1, Supplement1):65-66. Abstract Number S101.

23. Robinson KA, **Saldanha IJ**, McKoy NA. Setting priorities in clinical research: Identification of research needs from evidence-based guidelines. (*Poster*)
    23rd Annual North American Cystic Fibrosis Conference, October 15-17, 2009, Minneapolis, Minnesota, USA. *Pediatr Pulmonol* Suppl 2009; 32:385. Abstract Number 488.

**National**

24. **Saldanha IJ**, Wilson LM, McKoy NA, Bennett WL, Nicholson WK, Robinson KA. Identification of future research needs: methods of a pilot study. (*Poster*)
    2011 AHRQ Annual Conference, September 18-21, Bethesda, Maryland, USA

25. **Saldanha IJ**, Wilson LM, McKoy NA, Bennett WL, Nicholson WK, Robinson KA. Identification of future research needs: methods of a pilot study. (*Poster*)
    2011 Academy Health Annual Research Meeting, June 12-14, Seattle, Washington, USA

26. Robinson KA, **Saldanha IJ**, McKoy NA. Development of a framework for determining research gaps from systematic reviews. (*Poster*)
    2011 Academy Health Annual Research Meeting, June 12-14, Seattle, Washington, USA

**Local**

27. **Saldanha IJ**, Vedula SS, Yu T, Rosman L, Twose C, Li T, Dickersin K. Learning by doing - Teaching systematic review methods in 8 weeks. (*Poster*)
    2013 Institute for Excellence in Education Conference, April 5, Baltimore, Maryland, USA

28. Ervin A, Scherer R, Li T, Hawkins B, Lindsley K, Marrone M, Wang X, Vedula S, Yu T, **Saldanha I**, Dickersin K. The Cochrane Eyes and Vision Group US Project. (*Poster*)
    2012 Annual Wilmer Eye Institute Research Meeting, April 13, Baltimore, Maryland, USA

29. **Saldanha IJ**, Parikh R, Khan A, Desai S, Paranjape N. Space: Medicine beyond horizons. (*Oral*)
    2003 Bombay Medical Congress, March 1-2, Mumbai, India

## VOLUNTEER ACTIVITIES

2014–Present *Volunteer*
Our Daily Bread, Baltimore, Maryland, USA

2009–2013 *Volunteer*
Little Sisters of the Poor, St. Martin's Home, Catonsville, Maryland, USA

2005 *Volunteer*
School health camps, various locations, Maharashtra, India

2001 *Chief Organizer*
Gujarat earthquake donation efforts, Grant Medical College, Mumbai, India

## ADDITIONAL INFORMATION

### Date of Birth

October 29, 1982

### Place of Birth

Mumbai, India

### Personal Statement of Research Interests

My research expertise is drawn from my experience and rigorous training in epidemiology, public health, and medicine. I am interested in outcomes in research studies and, more broadly, in developing and improving methods for systematic reviews, randomized controlled trials, and epidemiologic studies. Most of my publications address methodology for comparative effectiveness research. My ongoing research involves selecting outcomes for clinical research using state-of-the-art methods.

### Personal Statement of Teaching Interests

My teaching experience covers the breadth of epidemiology, ranging from evidence generation (observational studies and clinical trials) to evidence synthesis (systematic reviews and meta-analyses) and policy (translational research). I am passionate about my research and teaching in the field of epidemiology.

### Keywords

Outcomes, epidemiology, randomized controlled trial, systematic review, meta-analysis, evidence-based healthcare, comparative effectiveness research, patient-centered outcomes research, reporting bias, publication bias, priority setting, cystic fibrosis, ophthalmology